



**University of
Zurich^{UZH}**

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2019

**Identifikation und Bereinigung von Inkonsistenzen in der Medizinischen
Statistik der Krankenhäuser (2000-16) und der Statistik der
Sozialmedizinischen Institutionen (2007-16)**

Zellweger, Ueli ; Bopp, Matthias

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-172587>

Monograph

Originally published at:

Zellweger, Ueli; Bopp, Matthias (2019). Identifikation und Bereinigung von Inkonsistenzen in der Medizinischen Statistik der Krankenhäuser (2000-16) und der Statistik der Sozialmedizinischen Institutionen (2007-16). Zürich: Universität Zürich, Institut für Epidemiologie, Biostatistik und Prävention.



**Universität
Zürich^{UZH}**

Institut für Epidemiologie, Biostatistik und Prävention

Identifikation und Bereinigung von Inkonsistenzen in der Medizinischen Statistik der Krankenhäuser (2000-16) und der Statistik der Sozialmedizinischen Institutionen (2007-16)

Zellweger U., Bopp M.

Korrekte Zitierweise:

Zellweger U., Bopp M.: Identifikation und Bereinigung von Inkonsistenzen in der Medizinischen Statistik der Krankenhäuser (2000-16) und der Statistik der Sozialmedizinischen Institutionen (2007-16)

Institut für Epidemiologie, Biostatistik und Prävention der Universität Zürich
Forschung und Dokumentation; Nr. 40 – Zürich 2019

Inhalt

1 Einleitung	1
Grundsätzliches Vorgehen.....	3
2 Inkonsistenzen in der MedStat	4
2.1 Fehlerhafte Anonyme Verbindungscode (AVC).....	4
2.1.1 Bereinigung systematischer Fehler nach Betrieb.....	5
2.1.2 Individuelle Bereinigung.....	7
2.2 Spitalaufenthalte über das Jahresende hinaus (B- und C-Fälle), für die es im Folgejahr keinen Fortsetzungsfall (A- oder C-Fall) gibt.....	10
2.3 Spitalaufenthalte im gleichen Betrieb mit identischem oder ähnlichem Eintritts- und Austrittsdatum: potentielle Mehrfachmeldungen	12
3 Die MedStat wieder konsistent machen	16
4 Konstellationen, die mit "vernünftigem" Aufwand nicht lösbar sind.....	18
4.1 B-/C-Fälle ohne Fortsetzungsrecord	18
4.2 "Mögliche" Mehrfachmeldungen	19
4.3 Komplexe Spitalaufenthalt-Serien	20
5 Inkonsistenzen in der SOMED	22
5.1 Fehlerhafter Anonymer Verbindungscode (AVC).....	22
5.2 Fehlende Fortsetzungsrecords im Folgejahr	22
5.3 Neue Variablen aus der Bereinigung von MedStat und SOMED	24
6 MedStat und SOMED kombiniert.....	25
6.1 Kennzeichnung überlappender Heim-/Spitalaufenthalte	25
6.2 Neue Variablen aus der Kombination von MedStat und SOMED	26
7 Weitere Datenfehler.....	27
7.1 Mehrfache Todesmeldungen für denselben AVC	27
7.2 Todesmeldung, aber trotzdem spätere Spital-/Heimaufenthalte.....	29
7.3 Verdächtig lange Aufenthaltsdauern (seit Geburt?)	30
8 Herkunft und Destination bei Spital- und Heimaufenthalten	32
8.1 Bewegung zu/von Spitälern	32
8.2 Bewegung zu/von Heimen.....	33
9 Fazit.....	35
9.1 Ungültige Fälle.....	35
9.2 Falsche Todesmeldung	37
9.3 Fehlerhafte Daten insgesamt	40
10 Liste der neu gebildeten Variablen und ihre Kennzahlen	41
Literatur	43

1 Einleitung

Die Medizinische Statistik der Krankenhäuser (MedStat) des Bundesamts für Statistik (BFS) ist mit jährlich 1.4-1.5 Mio. registrierten Spitalaufenthalten die umfangreichste Gesundheitsstatistik der Schweiz. Aus Datenschutzgründen wird die MedStat vom BFS in zwei sich gegenseitig ausschliessenden Versionen "GEO" und "TYPOL" angeboten [BFS 2018].

Eine grosse Stärke der MedStat liegt darin, dass Analysen nicht nur auf Ebene Fall (Hospitalisierung) sondern auch auf Ebene Patient durchgeführt werden können, d.h. Spitalaufenthalte desselben Individuums können auch dann zusammengeführt werden, wenn sie in unterschiedlichen Spitälern erfolgt sind. Möglich wird dies über einen nicht rückübersetzbaren Anonymen Verbindungscode (AVC), der bei der Datenerfassung in den Spitälern generiert und vom BFS in eine bei jeder Datenlieferung andere – aber innerhalb dieser Datenlieferung eindeutige – Personennummer umgewandelt wird (der Original-AVC wird vom BFS nicht an Dritte geliefert). In der Folge ist mit AVC jeweils die vom BFS generierte Personennummer und nicht der Original-AVC gemeint.

Der AVC wird mittlerweile nicht nur für die MedStat sondern auch für andere Bundesstatistiken zur Gesundheitsversorgung verwendet, nämlich seit 2007 für die Statistik der Sozialmedizinischen Institutionen (SOMED) und seit kurzem auch für die neuen Erhebungen im ambulanten Bereich [Creiser 2018]. Damit eröffnet sich ein grosses Potential für die Untersuchung von Verläufen auf Patientenebene ("pathways of care"), unabhängig davon, ob die Behandlung durch einen oder mehrere Leistungserbringer oder über mehrere Jahre verteilt erfolgt ist.

Die meisten uns bekannten bevölkerungsbezogenen (d.h. nicht nur ein Spital oder ein spezifisches Patientenkollektiv umfassenden) Analysen von Verläufen auf Patientenebene in der Schweiz stützen sich allerdings auf andere Datenquellen wie Krankenkassendaten [Reich et al 2013; Bähler et al 2016] oder das Schweizerische Haushalt-Panel [Weaver 2016]. Andere Arbeiten zu Verlaufstypen kombinierten die MedStat mit weiteren Datenquellen, aber ohne Verknüpfung auf Personenebene [Roth & Pellegrini 2018].

Analysen der MedStat auf Patienten- statt Fallebene waren bisher eher selten [z.B. aus dem Bereich der Psychiatrie Frick & Frick 2010] und meist auf Rehospitalisierungen und die Analyse von Abrechnungsstrukturen fokussiert [z.B. Widmer et al. 2011; Berlin et al. 2014]. Dabei wurde der konkrete Umgang mit Inkonsistenzen im AVC nur ausnahmsweise dokumentiert [Frick & Frick 2010]. Vereinzelt wurden auch Übertritte in Langzeitinstitutionen unter Verwendung des AVC untersucht [BFS 2009; Kohler et al. 2015; Widmer & Kohler 2015; Hedinger et al. 2015; Hedinger et al. 2016]. Eine kürzlich erschienene Publikation des Schweizerischen Gesundheitsobservatoriums stützte sich für die Definition von Lebensendverläufen (alle Übergänge zwischen Spital und Pflegeheim von in Institutionen Verstorbenen) ebenfalls auf den AVC statt die die unzuverlässigen Variablen "Aufenthalt vor dem Eintritt" und "Aufenthalt nach Austritt" ab [Füglister-Dousse & Pellegrini 2019]. Alle diese Arbeiten erwähnen Probleme mit fehlenden oder inkonsistenten AVC, verzichten aber auf eine genaue Dokumentation der Anzahl der Problemfälle und des Umgangs mit Inkonsistenzen im AVC oder in der Variable "Aufenthalt nach Austritt". Auch der Umgang mit komplexen Überlappungen (und den damit verbundenen potentiellen Mehrfachzählungen) wird nirgends thematisiert.

Im Rahmen des Nationalen Forschungsprojekts 74 "Gesundheitsversorgung" laufen mehrere Projekte¹, die den AVC nicht nur für eine Verbindung von MedStat und SOMED sondern auch für Verknüpfungen mit weiteren Datenquellen verwenden. Erfahrungen und Vorgehen bei Konsistenzproblemen wurden bisher noch keine publiziert. Eine systematische Evaluation des AVC scheint es bislang nicht zu geben und nur eine Arbeit [Frick & Frick 2010] hat exakt dokumentiert, wieviele AVC aus welchen Gründen von den Analysen ausgeschlossen wurden.

Die Datenqualität der MedStat hat sich seit dem ersten Registrationsjahrgang 1998 stark verbessert, doch fallen dem Benutzer doch immer wieder Inkonsistenzen (insbesondere beim AVC sowie bei den Variablen "Eintritt woher" und "Austritt nach") oder offensichtliche Mehrfachmeldungen auf [Zellweger & Bopp 2006, 2010, 2011, 2017]. Dabei zeigen sich Häufungen in bestimmten Jahren oder bei bestimmten Spitälern, deren gezielte Erkennung es möglich erscheinen lässt, einen Teil der Hospitalisierungen der betroffenen AVC für Verlaufsanalysen zu "retten". Mit den aktuell vom BFS gelieferten Standarddaten lassen sich solche "Problemfälle" ebenso wie überlappende Aufenthalte nicht mehr sauber erkennen. Auch die häufig systematischen Inkonsistenzen zwischen MedStat und SOMED bezüglich Aufenthaltsdauer und Sterbedatum [Zellweger & Bopp 2017] können mit der aktuellen Version der Standarddaten nicht mehr aufgelöst werden. Wir haben daher dem BFS angeboten, aufgrund unserer reichen Erfahrung eine umfassende Plausibilisierung der MedStat- und SOMED-Daten vorzunehmen und Zusatzvariablen zu generieren, die verschiedene Auswertungsoptionen der Standarddaten offen lassen. Dieser Bericht dokumentiert die vorgenommenen Plausibilisierungsarbeiten und die neu generierten Variablen.

Dabei haben wir uns mit dem BFS auf folgenden Ablauf geeinigt:

1. Das BFS liefert dem EBPI einen Spezialauszug der MedStat, Version TYPOL, mit exaktem Alter, Eintritts- und Austrittsdatum und anonymer Betriebsnummer, aber ohne Diagnosen und Behandlungen sowie dazu passende Angaben aus der SOMED. Das BFS speichert zwecks späterer Anreicherung der Originaldaten eine Linktabelle mit den originalen und den pseudonymisierten AVCs und Fallnummern.
2. Plausibilisierungsarbeiten am EBPI, Kennzeichnung problematischer Fälle und Bereinigung von Sequenznummern und Aufenthaltsdauern; Kennzeichnung, ob die letzte vorangegangene Hospitalisierung im gleichen Betrieb erfolgt ist.
3. Das EBPI liefert die Zusatzvariablen zusammen mit den erhaltenen pseudonymisierten AVCs und Fallnummern ans BFS.
4. Das BFS verknüpft diese Files mit der unter Abs. 1 erwähnten Linktabelle mit den Originaldaten.
5. Das EBPI löscht die unter Abs. 1 erwähnten Daten.
6. Das BFS fügt die vom EBPI generierten Zusatzinformationen an die MedStat- und SOMED-Datensätze (jeweils Version GEO) und liefert diese ans EBPI.

¹ z.B. "Soziale Ungleichheiten in der stationären Gesundheitsversorgung in der Schweiz" (L. Bayer-Oglesby) oder "Bessere Daten zur Qualität der häuslichen Pflege (Spitex)" (P. Rüesch)

Grundsätzliches Vorgehen

In einem ersten Schritt werden Inkonsistenzen innerhalb der einzelnen Datensätze (MedStat bzw. SOMED) gesucht und bereinigt; in einem weiteren Schritt Inkonsistenzen, die beim Zusammenführen der beiden Datensätze auftreten und sekundäre Inkonsistenzen, die manchmal notwendigerweise durch die Bereinigung der primären Inkonsistenzen entstehen. (Letztere entstehen beispielsweise dadurch, dass durch die Ungültigkeitsdeklaration eines Records mit fehlerhaftem anonymem Verbindungscode oder bei Doppelmeldungen die nachfolgenden Sequenznummern nicht mehr korrekt sind).

Neben den Inkonsistenzen innerhalb der MedStat und der SOMED (Kapitel 2 bis 5) werden durch die Kombination beider Datenquellen zusätzliche Angaben auf ihre Konsistenz geprüft, z.B. Überlappungen von Spital- mit Heimaufenthalten, mehrfache Todesmeldungen für denselben AVC, Spital- und/oder Heimaufenthalte nach einer Todesmeldung und verdächtig lange Aufenthaltsdauern. (Kapitel 6 und 7). Schliesslich werden die dabei gewonnenen Resultate für eine Harmonisierung der Angaben zu den individuellen Patientenströmen (Herkunft und Destination bei Spital-/Heimaufenthalten) genutzt (Kapitel 8).

2 Inkonsistenzen in der MedStat

In drei Bereichen gibt es in grösserem Ausmass Inkonsistenzen, deren Bereinigung im Folgenden dokumentiert wird:

- (A) Fehlerhafte Anonyme VerbindungsCodes (AVC)
- (B) Spitalaufenthalte über das Jahresende hinaus (B- und C-Fälle), für die es im Folgejahr keinen Fortsetzungsfall (A- oder C-Fall) gibt
- (C) Spitalaufenthalte im gleichen Betrieb mit identischem oder ähnlichem Eintritts- und Austrittsdatum: Duplikate oder multiple Aufenthaltsmeldungen.

2.1 Fehlerhafte Anonyme VerbindungsCodes (AVC)

Der AVC der MedStat soll die Zuordnung verschiedener Spitalaufenthalte einer Person ermöglichen. Wenn es sich wirklich um dieselbe Person handelt, sollten deshalb die demografischen Angaben übereinstimmen. Die MedStat enthält drei Variablen mit demografischen Angaben zur Person:

1. Alter in (vollendeten) Jahren beim Spitaleintritt
2. Geschlecht
3. Wohnregion.

ad 1) Aus dem Eintrittsalter lässt sich das Geburtsjahr auf ein Jahr genau errechnen.

ad 2) Das Geschlecht müsste bei jedem Spitalaufenthalt dasselbe sein.

ad 3) Die Wohnregion muss nicht identisch sein, bei einem Umzug kann diese ändern. Gänzlich unterschiedliche Wohnregionen können aber zumindest einen Hinweis geben / Verdacht wecken, dass es sich nicht um verschiedene Spitalaufenthalte derselben Person handelt sondern um unterschiedliche Personen. (Aus Datenschutzgründen stand uns die Wohnregion für dieses Projekt nicht zur Verfügung).

Für die Jahre 2000 bis 2016 enthält unser Datenauszug insgesamt 23'788'383 Datensätze (Hospitalisierungen), davon sind

- 23'375'696 A-Fälle (Spitalaustritt im Statistikjahr),
- 389'817 B-Fälle (Eintritt, aber kein Austritt im Statistikjahr, also Sylvester/Neujahr im Spital)
- 22'870 C-Fälle (schon vor dem Statistikjahr eingetreten und noch nicht wieder ausgetreten).

513'467 Hospitalisierungen (2.2%) hatten einen auf den ersten Blick plausiblen AVC (d.h., grösser als 0; insgesamt 70'302 unterschiedliche AVCs), unter dem aber entweder beide Geschlechter oder Geburtsjahre mit einer Differenz grösser als 1 registriert worden sind (Tab. 1).

Tab. 1 Hospitalisierungen von AVCs mit inkonsistenten Angaben zu Alter und Geschlecht, nach Statistikjahr

Jahr	Hospitalisierungen	Jahr	Hospitalisierungen
2000	16'851	2009	27'970
2001	23'177	2010	24'674
2002	23'285	2011	40'526
2003	26'803	2012	20'197
2004	31'768	2013	18'977
2005	69'012	2014	17'826
2006	40'611	2015	16'932
2007	36'405	2016	17'422
2008	61'031	Total	513'467

Die nähere Durchsicht dieser halben Million Hospitalisierungen zeigt:

1. In sehr vielen Fällen ist es eine einzelne Hospitalisierung, die nicht in eine ansonsten konsistente Reihe von Hospitalisierungen unter diesem AVC passt;
2. Meistens treten AVCs mit inkonsistenten Angaben zu Alter und Geschlecht gehäuft auf, und zwar bei einigen wenigen Spitälern in einzelnen Jahren, was auf einen systematischen, nicht-zufälligen Fehler hindeutet.

Lösungsansatz:

Wenn es so ist, dass eine einzelne oder zwei fehlerhafte Registrierungen eine ansonsten konsistente Reihe von Hospitalisierungen einer Person korrumpiert, wäre es schade, diese Reihe als Ganzes für inkonsistent und damit nicht auswertbar zu deklarieren. Vielmehr kann durch die "Elimination" (bzw. durch Ungültigerklärung) von Querschlägern der Grossteil dieser Reihe "gerettet" werden. Damit kann die Datenqualität deutlich verbessert (konsistenter) gemacht werden.

Nachstehend werden die systematischen Fehler und deren Behebung beschrieben. Zur Kennzeichnung werden die Datensätze mit einer Zusatzvariablen versehen:

INV_VC Ungültiger Anonymer Verbindungscode (AVC)

- | | |
|---|--|
| 0 | nein (AVC ist unverdächtig) |
| 1 | systematischer Fehler in einem Betrieb |
| 2 | einzelne Hospitalisierungen, die nicht zur Mehrheit der übrigen Hospitalisierungen unter diesem AVC passen |
| 3 | AVC umfasst nur wenige Hospitalisierungen, deshalb kann nicht entschieden werden, welche Hospitalisierungen vermutlich nicht zu diesem AVC gehören |

2.1.1 Bereinigung systematischer Fehler nach Betrieb

Eine systematische Durchsicht der (anonymisierten) Spitalnummern, die von solchen Inkonsistenzen betroffen sind, zeigt auffällige Häufungen nach bestimmten Betrieben und Jahren. Dabei ist primär nicht festzustellen, ob ein Betrieb diese Fehler selber produziert hat, oder ob er umgekehrt "Opfer" eines solchen Fehlers in einem anderen Betrieb ist. Eine stichprobenweise Durchsicht erlaubt aber in der Regel plausibel festzulegen, wo der Fehler passiert ist. Allerdings ist dies nicht immer einfach, besonders, wenn zwei Betriebe "irgendwie" miteinander verbunden sind (d.h., wenn häufig Patientenflüsse zwischen ihnen stattfinden; z.B. zuerst Akutsomatik in Betrieb A und dann Reha in Betrieb B). In solchen Fällen müssen die entsprechenden Datensätze miteinander verglichen werden, damit dann aufgrund der Mehrheitsverhältnisse entschieden werden kann, in welchem Betrieb systematisch ein falscher AVC vergeben wurde. Gemäss unseren Überprüfungen war dies in 28 Betrieben zumindest während eines Jahres der Fall (Tab. 2).

Von den 513'467 Hospitalisierungen mit inkonsistentem Alter und/oder Geschlecht konnten 107'511 als systematische Fehler bei der Vergabe des AVC in einem dieser 28 Betriebe eingestuft werden. Mit deren Ausschluss konnten 401'541 Hospitalisierungen bereinigt werden, d.h. die verbleibenden Hospitalisierungen mit dem entsprechenden AVC waren nun bezüglich Alter und Geschlecht konsistent. Nach dieser Bereinigung verblieben 4'415 Hospitalisierungen von 939 AVC mit Inkonsistenzen bezüglich Alter und/oder Geschlecht. Diese wurden in der Folge individuell evaluiert.

Tab. 2 Anzahl Hospitalisierungen mit inkonsistentem Alter/Geschlecht in 28 Betrieben mit systematischen Fehlern beim AVC, nach Kalenderjahren

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Subtotal
Hospitalisierungen mit fehlerhaftem AVC	16851	23177	23285	26803	31768	69012	40611	36405	61031	27970	24674	40526	20197	18977	17826	16932	17422	513'467
Anonyme Betriebs-Nr.																		
SPI1	0	0	0	0	66	68	71	0	0	0	0	0	0	0	0	0	0	205
SPI2	0	1324	0	0	0	0	0	0	0	0	0	212	0	0	0	0	0	1'536
SPI3	0	0	0	59	61	0	0	0	0	0	0	0	0	0	0	0	0	120
SPI4	0	0	0	0	0	0	0	244	0	0	0	0	0	0	0	0	0	244
SPI5	0	0	0	434	0	0	0	0	0	0	0	0	0	0	0	0	0	434
SPI6	0	0	0	0	0	0	0	0	0	0	0	18530	0	0	0	0	0	18'530
SPI7	0	0	0	0	0	0	0	0	6501	0	0	0	0	0	0	0	0	6'501
SPI8	0	0	170	0	0	37996	0	0	0	0	0	0	0	0	0	0	0	38'166
SPI9	0	0	0	0	0	0	0	363	0	0	0	0	0	0	0	0	0	363
SPI10	0	0	0	0	0	0	2494	2382	2194	2325	0	0	0	0	0	0	0	9'395
SPI11	0	0	0	0	0	0	0	0	8966	0	0	953	0	0	0	0	0	9'919
SPI12	0	2017	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2'017
SPI13	0	0	0	0	0	0	0	0	0	78	0	0	0	0	0	0	0	78
SPI14	0	0	0	0	0	0	0	0	7478	0	0	0	0	0	0	0	0	7'478
SPI15	0	0	0	0	0	0	0	0	2302	0	0	0	0	0	0	0	0	2'302
SPI16	0	0	0	0	0	0	0	36	0	0	0	0	0	0	0	0	0	36
SPI17	0	0	0	0	49	0	0	0	0	0	0	0	0	0	0	0	0	49
SPI18	0	0	0	46	0	0	0	63	0	0	0	0	0	0	0	0	0	109
SPI19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	20
SPI20	0	65	64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	129
SPI21	0	0	0	0	0	0	0	0	2480	0	0	0	0	0	0	0	0	2'480
SPI22	2718	0	0	0	0	0	0	0	0	0	0	224	0	0	0	0	0	2'942
SPI23	0	0	0	0	0	0	0	0	54	0	0	0	0	0	0	0	0	54
SPI24	0	0	0	0	0	0	0	417	0	0	0	0	0	0	0	0	0	417
SPI25	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	50
SPI26	0	0	0	120	41	0	0	0	0	0	0	0	0	0	0	0	0	161
SPI27	0	0	0	100	99	114	0	0	0	0	0	0	0	0	0	0	0	313
SPI28	0	986	956	0	740	699	0	0	0	0	0	82	0	0	0	0	0	3'463
mit fehlerhaftem AVC in diesen Betrieben	2718	4392	1190	759	1056	38877	2565	3555	29'975	2403	0	20'001	0	0	0	20	0	107'511
bereinigte	14'057	18'584	21'829	25'749	30'493	29'887	37'762	32'614	30'849	25'371	24'441	19'250	19'996	18'858	17'725	16'781	17'295	401'541
ungelöste Inkonsistenzen	76	201	266	295	219	248	284	236	207	196	233	1275	201	119	101	131	127	4'415

2.1.2 Individuelle Bereinigung

Bei 182 AVCs mit insgesamt 1'033 Hospitalisierungen wurde der Verbindungscode bei 227 Hospitalisierungen als ungültig gewertet (dabei war oft der AVC von Mutter und Kind bei der Geburt derselbe) bzw. es konnten 806 Hospitalisierungen "gerettet" werden.

Bei weiteren 53 AVCs mit zusammen 106 Hospitalisierungen war klar, dass es sich um unterschiedliche Personen handeln muss, aber es konnte nicht festgelegt werden, bei welcher Hospitalisierung der AVC als falsch und damit ungültig gewertet werden sollte.

Bei den übrigen 704 AVCs mit insgesamt 3'329 Hospitalisierungen kamen wir zum Schluss, dass es sich um dieselbe Person handelt, mit einem Datenfehler bei Alter oder Geschlecht (433 AVCs mit insgesamt 1'185 Hospitalisierungen betrafen einen einzigen Betrieb, wo in einem Jahr das Geburtsjahr um immer genau 2 Jahre differierte).

Bei dieser individuellen Bereinigung konnten somit insgesamt 4'135 Hospitalisierungen von 886 AVCs konsistent gemacht werden.

Durch die Bereinigung in den Arbeitsschritten A1 und A2 wurde der AVC von 107'791 (107'511+227+53) Hospitalisierungen als ungültig erklärt, wodurch 405'676 (401'541+4'135) Hospitalisierungen konsistent gemacht werden konnten.

Wenn in einem bestimmten Betrieb in einem bestimmten Jahr eine grosse Mehrheit der AVCs als systematisch ungültig eingestuft wird, liegt der Gedanke nahe, dass der AVC auch bei den auf den ersten Blick unauffälligen Hospitalisierungen falsch ist (dass also Alter und Geschlecht nur zufällig mit den anderen Hospitalisierungen übereinstimmen).

Je nach konkreter Fragestellung/Auswertungsstrategie kann es deshalb Sinn machen, alle Hospitalisierungen dieses Betriebs im entsprechenden Jahr für ungültig zu erklären (z.B. bis zur Zeile 21 oder auch 25 in Tab. 3) – oder umgekehrt: Ab Zeile 32 (oder schon 27) könnte man auch davon ausgehen dass es sich um einen korrekten AVC mit falscher Alters- bzw. Geschlechtsangabe handelt .

Um das bei künftigen Analysen flexibel handhaben zu können, wurde für jede Hospitalisierung der Prozentwert der in diesem Jahr in diesem Betrieb als systematisch ungültig erachteten Hospitalisierungen festgehalten (Tab. 3).

<i>INV_VC_PERC Anteil systematisch falscher AVCs in diesem Betrieb und Statistikjahr</i> Wertebereich 0.0 - 1.0 (entspricht 100%)

Tab. 3a Anteil Hospitalisierungen mit systematisch falschem Verbindungscode (AVC) nach Betrieb und Jahr

Betrieb/Jahr		Total Hospitalisierungen	AVC mit nur einer Hospitalisierung	AVC mit mehr als einer Hospitalisierung im entsprechenden Jahr			
				Total Hospitalisierungen	N Fehler (siehe Tab. 2)	% als systematisch ungültig erachtet	verbleibende Hospitalisierungen
1	SPI10 / 2007	2'382	0	2'382	2'382	100.0%	0
2	SPI10 / 2009	2'325	0	2'325	2'325	100.0%	0
3	SPI10 / 2008	2'194	0	2'194	2'194	100.0%	0
4	SPI24 / 2007	417	0	417	417	100.0%	0
5	SPI9 / 2007	363	0	363	363	100.0%	0
6	SPI4 / 2007	244	0	244	244	100.0%	0
7	SPI18 / 2007	63	0	63	63	100.0%	0
8	SPI25 / 2007	50	0	50	50	100.0%	0
9	SPI16 / 2007	36	0	36	36	100.0%	0
10	SPI19 / 2015	79	66	20	20	100.0%	0
11	SPI10 / 2006	2'496	0	2'496	2'494	99.9%	2
12	SPI6 / 2011	21'579	2'883	18'696	18'530	99.1%	166
13	SPI28 / 2005	706	0	706	699	99.0%	7
14	SPI21 / 2008	2'513	0	2'513	2'480	98.7%	33
15	SPI22 / 2000	3'291	537	2'754	2'718	98.7%	36
16	SPI2 / 2001	1'422	72	1'350	1'324	98.1%	26
17	SPI18 / 2003	47	0	47	46	97.9%	1
18	SPI28 / 2002	978	0	978	956	97.8%	22
19	SPI28 / 2001	1'039	3	1'036	986	95.2%	50
20	SPI8 / 2005	44'858	4'708	40'150	37'996	94.6%	2'154
21	SPI28 / 2004	814	5	809	740	91.5%	69
22	SPI11 / 2008	12'621	2'543	10'078	8'966	89.0%	1'112
23	SPI7 / 2008	9'266	1'622	7'644	6'501	85.0%	1'143
24	SPI14 / 2008	11'316	2'002	9'314	7'478	80.3%	1'836
25	SPI12 / 2001	2'992	162	2'830	2'017	71.3%	813
26	SPI13 / 2009	164	8	156	78	50.0%	78
27	SPI5 / 2003	3'059	568	2'491	434	17.4%	2'057
28	SPI2 / 2011	1'359	35	1'324	212	16.0%	1'112
29	SPI28 / 2011	700	18	682	82	12.0%	600
30	SPI11 / 2011	10'526	2'152	8'374	953	11.4%	7'421
31	SPI15 / 2008	24'892	2'679	22'213	2'302	10.4%	19'911
32	SPI22 / 2011	5'300	473	4'827	224	4.6%	4'603
33	SPI26 / 2003	3'627	599	3'028	120	4.0%	2'908
34	SPI26 / 2004	2'594	631	1'963	41	2.1%	1'922
35	SPI27 / 2003	8'391	1'595	6'796	100	1.5%	6'696
36	SPI20 / 2001	8'333	3'520	4'813	65	1.4%	4'748
37	SPI27 / 2004	9'424	1'657	7'767	99	1.3%	7'668
38	SPI20 / 2002	8'601	3'827	4'774	64	1.3%	4'710
39	SPI3 / 2003	5'557	503	5'054	59	1.2%	4'995
40	SPI3 / 2004	5'562	453	5'109	61	1.2%	5'048
41	SPI27 / 2005	11'147	1'945	9'202	114	1.2%	9'088
42	SPI1 / 2006	12'556	3'578	8'978	71	0.8%	8'907
43	SPI1 / 2004	11'562	2'606	8'956	66	0.7%	8'890
44	SPI1 / 2005	12'603	2'670	9'933	68	0.7%	9'865
45	SPI8 / 2002	35'788	1'622	30'252	170	0.6%	30'082
46	SPI17 / 2004	10'765	2'058	8'707	49	0.6%	8'658
47	SPI23 / 2008	58'156	75	58'080	54	0.1%	58'026
systematischer Fehler		374'757	47'875	322'974	107'511	(33.3%)	215'463
Hospitalisierungen Total		23'788'383	4'618'276	19'170'107	107'511	0.6%	19'062'596

Zu zwei Betrieben je eine Bemerkung:

- Im Spital 13 (Tab 3a: Zeile 26), wo genau die Hälfte aller AVCs mit mehrfachen Auf-
enthalten falsch sind, scheint es sich oft nicht wirklich um systematische Fehler zu handeln,
sondern eher um eine Häufung von Schreib-/Rechnungsfehlern (z.B. Jahrgang 1917 statt
1967 und damit Alter 92 statt 42; generell: ähnliche Zahlen wie 5 oder 3 statt 8; 7 statt 2).
- Im Spital 28 (Tab 3a: Zeilen 13, 18, 19, 21) sind die Jahre 2001 bis 2005 (mit Ausnahme
von 2003) systematisch fehlerhaft. Es fällt auf, dass in diesen vier Jahren fast kein AVC nur
ein einziges Mal vorkommt – dies im Gegensatz zum Jahr 2003 mit sehr vielen AVCs mit
nur einer Hospitalisierung (die naturgemäss nicht als systematisch ungültig erkannt werden
können). (Fast) jeden entsprechenden AVC der Jahre 2002 und 2005 findet man auch im
jeweiligen Vorjahr (Tab. 3b), wobei Geburtsjahr und Geschlecht auf keinen systematischen
Fehler hindeuten. Man gewinnt den Eindruck, die AVCs des jeweiligen Vorjahres seien
irgendwie in das folgende Jahr "hineingerutscht".

Tab. 3b AVCs in Spital 28 nach Jahr (Auszug)

AVC	2001	2002	2003	2004	2005
a			2		
b				1	1
c			1		
d	3	2			
e	4	4			
f			1		
g			1		
h				1	1
i			1	1	1
J			2		
k			1		
l	1	1			
m			1		
n			1		
o			1		
p				1	1
q	1	1			
r	1	1			
s	1	1			
t	1	1		1	1
u	1	1			
v			1		
w				1	1
x			1		
y			1		
z			1	1	1
aa			1		
ab			1		
ac				1	1
ad			1		
ae	3	3		2	2

AVC	2001	2002	2003	2004	2005
af	1	1			
ag	1	1			
ah			1		
ai			1		
aJ			2		
ak	1	1			
al			3		
am				1	1
an	1	1			
ao			1		
ap	1	1			
aq				1	
ar			1		
as			1		
at			4		
au			1		
av				1	1
aw			1		
ax			1		
ay			1		
az			1		
ba			1		
bb				1	1
bc				1	1
bd			1		
be			1		
bf	2	2			
bg	1	1		1	1
bh				1	
bi	1	1			
bj			1		

AVC	2001	2002	2003	2004	2005
bk				1	1
bl	1	1	1		
bm	1	1			
bn	1	1			
bo			1		
bp				1	
bq	1	1			
br			1		
bs			1		
bt			2		
bu	1	1			
bv				1	1
bw			1		
bx				1	1
by	1	1			
bz				1	1
ca	1	1			
cb			1		
cc				3	3
cd	1				
ce	3	3		1	1
cf			2		
cg			1		
ch			1		
ci				1	1
cJ			1		
ck	1	1			
cl			2		
cm			1		
cn	3	2	1	2	2
co	1	1			

2.2 Spitalaufenthalte über das Jahresende hinaus (B- und C-Fälle), für die es im Folgejahr keinen Fortsetzungsfall (A- oder C-Fall) gibt

Für longitudinale Analysen auf Patientenebene sind B-Fälle, für die es im Folgejahr einen A- oder C-Fall gibt, redundant, weil erstens bei B-Fällen keine Diagnosen und Behandlungen angeführt sind und zweitens die wenigen übrigen Informationen dieses B-Falls in einem A- (oder C-)Fall des Folgejahres ebenfalls enthalten sein müssten.

Für C-Fälle mit Fortsetzungsfall im Folgejahr gilt dies eigentlich auch, allerdings werden nicht immer alle Diagnosen und Behandlungen im Fortsetzungsjahr aufgenommen.

Für die B- und C-Fälle der Jahre 2000-2015 mit gültigem AVC (Tab. 4) müsste es jeweils im Folgejahr einen A- oder C- Fall geben (die 2016er Fälle könnten nur unter Beizug des für uns nicht verfügbaren Statistikjahrgangs 2017 evaluiert werden).

Tab. 4 B- und C-Fälle mit gültigem AVC, nach Statistikjahr

	Jahr	Adm. Meldung mit reduziertem Datensatz (B-Fall)	Adm. Meldung mit vollständigem Datensatz (C- Fall)	Total
fehlender AVC*	2000-2002	5'826	1'629	7'455
	2003-2008	136	6	142
	2009-2016	30	4	34
als ungültig erachteter AVC**	2000-2002	261	8	269
	2003-2008	876	0	876
	2009-2016	153	3	156
	Total	7'282	1'650	8'932
gültiger AVC	2000	17'756	1'687	19'443
	2001	20'831	1'705	22'536
	2002	21'869	2'247	24'116
	2003	24'037	2'550	26'587
	2004	22'765	2'602	25'367
	2005	22'224	2'045	24'269
	2006	23'781	1'769	25'550
	2007	22'387	1'472	23'859
	2008	22'185	1'614	23'799
	2009	21'423	974	22'397
	2010	21'877	486	22'363
	2011	21'321	448	21'769
	2012	23'119	348	23'467
	2013	23'642	373	24'015
	2014	23'710	290	24'000
	2015	24'105	308	24'413
	2000-2015	357'032	20'918	377'950
	2016	25'503	302	25'805
	Total	382'535	21'220	403'755
Total		389'817	22'870	412'687

* kein AVC vorhanden in der MedStat

** als fehlerhaft / ungültig erachteter AVC gemäss Kapitel 2.1

Längst nicht alle der 377'950 B- und C-Fälle der Jahre 2000-2015 mit gültigem AVC haben im Folgejahr einen A- oder C-Fall als Fortsetzung (Tab. 5a+b).

Idealfall ist selbes Eintrittsdatum und selbe Spitalnummer wie der B-/C-Fall im Vorjahr.

Möglicherweise auch korrekt ist eine andere Spitalnummer, aber identisches Eintrittsdatum oder Eintritt 1. Januar im Folgejahr, z.B. weil Spitäler aufgehoben worden sind oder fusioniert haben.

Tab. 5a Fortsetzungshospitalisierungen von B-Fällen, nach Jahr

Jahr	identische Spitalnummer und Eintrittsdatum	andere Spitalnummer, identisches Eintrittsdatum	identische Spitalnummer, Eintrittsdatum 1. Januar	keine Folgehospitalisierung	Total
2000	14'128	99	134	3'395	17'756
2001	16'169	719	566	3'377	20'831
2002	18'524	299	166	2'880	21'869
2003	20'457	116	117	3'347	24'037
2004	18'203	440	514	3'608	22'765
2005	19'107	493	196	2'428	22'224
2006	18'764	1'252	81	3'684	23'781
2007	19'185	37	15	3'150	22'387
2008	18'689	21	11	3'464	22'185
2009	15'609	3'571	100	2'143	21'423
2010	19'793	96	9	1'979	21'877
2011	16'814	541	8	3'958	21'321
2012	19'556	125	13	3'425	23'119
2013	21'188	169	12	2'273	23'642
2014	19'861	62	154	3'633	23'710
2015	21'739	93	6	2'267	24'105
2000-2015	297'786	8'133	2'102	49'011	357'032
	83%	2%	1%	14%	100%
2016	0	0	0	25'503	25'503
Total	297'786	8'133	2'102	74'514	382'535

Tab. 5b Fortsetzungshospitalisierungen von C-Fällen, nach Jahr

Jahr	identische Spitalnummer und Eintrittsdatum	andere Spitalnummer, identisches Eintrittsdatum	identische Spitalnummer, Eintrittsdatum 1. Januar	keine Folgehospitalisierung	Total
2000	1'303	6	0	378	1'687
2001	1'006	312	1	386	1'705
2002	1'792	1	27	427	2'247
2003	1'937	3	0	610	2'550
2004	1'466	137	5	994	2'602
2005	1'562	39	2	442	2'045
2006	1'279	72	2	416	1'769
2007	1'210	0	1	261	1'472
2008	723	1	0	890	1'614
2009	528	53	0	393	974
2010	442	2	0	42	486
2011	300	17	0	131	448
2012	275	0	0	73	348
2013	265	0	0	108	373
2014	224	0	0	66	290
2015	279	0	0	29	308
2000-2015	14'591	643	38	5'646	20'918
	70%	3%	0%	27%	100%
2016	0	0	0	302	302
Total	14'591	643	38	5'948	21'220

FOLGEHOSP Fortsetzungshospitalisierung prospektiv (die Variable steht auf jedem B-/C-Record)

- 1 (kein B-/C-Record)
- 0 keine Folgehospitalisierung
- 1 identisches Spital und Eintrittsdatum
- 2 anderes Spital aber identisches Eintrittsdatum
- 3 identisches Spital, aber Eintrittsdatum 1. Januar

Wenn das Spital (mit derselben **oder mit einer neuen** Betriebsnummer) einen Fortsetzungs-record mit Eintrittsdatum 1. Januar erstellt hat, muss die Aufenthaltsdauer dieses Fortsetzungs-falls entsprechend angepasst werden.

Tab. 6 Zusammenfassung der Tabellen 5a und 5b

Fortsetzungshospitalisierung		Anonymer Verbindungscode (AVC)					Total
		fehlend	gültig		ungültig		
				systematischer Fehler in Betrieb	einzelne Hospitalisierung, die nicht passt	AVC falsch, aber unklar, welche Hospitalisierung	
B-/C-Fälle	keine	7'631	80'462	1'286	10	5	89'394
	identisches Spital / Eintrittsdatum		312'377	0	0	0	312'377
	anderes Spital /identisches Eintrittsdatum		8'776	0	0	0	8'776
	identisches Spital, Eintrittsdatum 1.Januar		2'140	0	0	0	2'140
A-Fälle		130'964	23'138'189	106'225	217	101	23'375'696
Total		138'595	23'541'944	107'511	227	106	23'788'383

2.3 Spitalaufenthalte im gleichen Betrieb mit identischem oder ähnlichem Eintritts- und Austrittsdatum: potentielle Mehrfachmeldungen

Von den insgesamt 23'788'383 Hospitalisierungen 2000-2016 haben 138'595 keinen AVC und 107'844 einen von uns als ungültig markierten AVC (Tab. 6). Von den verbliebenen 23'514'944 Hospitalisierungen sind 4'479'681 AVCs (19%) singulär, d.h. es gibt in diesem Jahr keine weiteren Hospitalisierungen mit diesem AVC; bei 19'062'263 Hospitalisierungen kommt der entsprechende AVC mehr als einmal vor, d.h. es kann davon ausgegangen werden, dass die entsprechende Person mehrfach hospitalisiert worden ist (Tab. 7a).

Tab. 7a Ein- und mehrfache Hospitalisierungen

Jahr	Total	Ohne oder ungültiger AVC	gültiger AVC	"nur" eine Hospitalisierung	mehr als eine Hospitalisierung N	%
2000	1'184'229	85'612	1'098'617	262'547	836'070	76.1%
2001	1'302'964	40'388	1'262'576	253'621	1'008'955	79.9%
2002	1'346'129	18'617	1'327'512	244'726	1'082'786	81.6%
2003	1'390'903	929	1'389'974	244'702	1'145'272	82.4%
2004	1'409'743	1'218	1'408'525	248'850	1'159'675	82.3%
2005	1'466'467	39'092	1'427'375	246'148	1'181'227	82.8%
2006	1'500'664	2'725	1'497'939	239'522	1'258'417	84.0%
2007	1'538'323	3'699	1'534'624	239'032	1'295'592	84.4%
2008	1'584'754	31'237	1'553'517	266'420	1'287'097	82.9%
2009	1'328'127	2'482	1'325'645	212'836	1'112'809	83.9%
2010	1'345'245	72	1'345'173	248'361	1'096'812	81.5%
2011	1'363'698	20'057	1'343'641	246'458	1'097'183	81.7%
2012	1'353'521	185	1'353'336	292'844	1'060'492	78.4%
2013	1'374'454	27	1'374'427	276'028	1'098'399	79.9%
2014	1'401'014	18	1'400'996	303'408	1'097'588	78.3%
2015	1'430'201	53	1'430'148	331'364	1'098'784	76.8%
2016	1'467'947	28	1'467'919	322'814	1'145'105	78.0%
Total	23'788'383	246'439	23'541'944	4'479'681	19'062'263	81.0%

Insgesamt besteht bei rund 50'000 dieser 19 Millionen Hospitalisierungen (gut ein Viertel Prozent) der Verdacht, dass es sich um multiple Aufenthaltsmeldungen handelt, die sich in Wirklichkeit auf nur rund 25'000 Spitalaufenthalte beziehen. Bei der Identifikation von möglichen multiplen Aufenthaltsmeldungen wurde immer vorausgesetzt, dass die Betriebsnummer aller entsprechenden Hospitalisierungen jeweils identisch war.

Es ist nicht immer eindeutig feststellbar, ob es sich bei Spitalaufenthalten mit überschneidenden oder sogar sich ganz überdeckenden Ein- bzw. Austritts-Daten um eine Verlegung des Patienten (aber mit verspäteter Austrittsmeldung der einen Abteilung), oder um eine doppelte administrative Meldung handelt. Es wurde deshalb eine Hierarchisierung nach Wahrscheinlichkeit einer Doppelmeldung vorgenommen und in der Variable DOPPEL wie folgt codiert (wobei sich jeweils die tiefere Zahl des Zahlenpaares auf die "erste" Meldung, die zweite auf die "weitere/n" Meldung/en bezieht).

DOPPEL Potentielle Mehrfachmeldung		
	Definition	Bezeichnung im Text
0	keine Mehrfachmeldung	
1 & 2	exakte Übereinstimmung	"sichere" Mehrfachmeldung
3 & 4	selbe Eintrittsstunde, selbes Austrittsdatum (ohne Stunde)	"wahrscheinliche" Mehrfachmeldung
5 & 6	selbes Eintrittsdatum (ohne Stunde), selbe Austrittsstunde	"wahrscheinliche" Mehrfachmeldung
7 & 8	selbes Eintrittsdatum, die kürzere Hospitalisierung ist aber maximal 1 Tag lang	"mögliche" Mehrfachmeldung
9 & 10	selbes Eintrittsdatum, aber beide Hospitalisierungen sind länger als 1 Tag	"eher keine" Mehrfachmeldung
11	verworfen, d.h. keinesfalls als Mehrfachmeldung zu betrachten	

Bei 12'735 AVCs mit insgesamt 27'528 Hospitalisierungen sind bei mindestens zwei Hospitalisierungen sowohl Eintrittsdatum und Austrittsdatum (jeweils inkl. Stunde) als auch Statistikfall (A-, B-, C-Fall) identisch (Tab. 7b). Es kann somit von eindeutigen bzw. "sicheren" Mehrfachmeldungen gesprochen werden.

Beispiel:

jahr	betrieb	seq_u	statf	eintritt	ein_s	austritt	aus_s	born_y	doppel	v14_01
2004	SPI1	1	A	03-AUG-2004	19	04-AUG-2004	8	1999	0	M200
2006	SPI1	.	A	12-JUL-2006	20	13-JUL-2006	11	1999	1	M200
2006	SPI1	2	A	12-JUL-2006	20	13-JUL-2006	11	1999	2	M200
2011	SPI2	3	A	25-FEB-2011	13	26-FEB-2011	15	1999	0	M200

Tab. 7b Anzahl AVCs mit mehr als einer Hospitalisierungen: "Sichere" Mehrfachmeldungen

Jahr	"erste" Hospitalisierung	weitere Hospitalisierungen	Total
2000	1'169	1'276	2'445
2001	727	750	1'477
2002	2'300	2'316	4'616
2003	4'860	5'194	10'054
2004	413	418	831
2005	346	359	705
2006	1'102	1'655	2'757
2007	280	280	560
2008	688	1'545	2'233
2009	186	293	479
2010	136	136	272
2011	108	138	246
2012	66	75	141
2013	208	208	416
2014	56	57	113
2015	27	27	54
2016	63	66	129
Total	12'735	14'793	27'528

Neben diesen "sicheren" gibt es *"wahrscheinliche"* Mehrfachmeldungen, bei denen bloss Eintritts- oder Austrittsstunde nicht übereinstimmen (häufig ist bei der einen Meldung die Eintritts- respektive Austrittsstunde Null). Man bekommt den Eindruck, dass es sich um Patienten handeln könnte, die auf eine andere Abteilung verlegt und dort als Neueintritt noch einmal erfasst worden sind, während die ursprüngliche Abteilung als Austrittsdatum nicht den Verlegungszeitpunkt, sondern erst den endgültigen Spitalaustritt gemeldet hat.

Beispiele:

jahr	betrieb	seq_u	statf	eintritt	ein_s	austritt	aus_s	born_y	doppel	v14_01
2001	SPI1	.	A	08-NOV-2001		08-NOV-2001		1930	0	
2001	SPI1	.	A	15-NOV-2001	0	15-NOV-2001	15	1930	0	M600
2003	SPI2	.	A	26-MAY-2003	10	04-JUN-2003	0	1931	4	M200
2003	SPI2	1	A	26-MAY-2003	10	04-JUN-2003	11	1931	3	M200
2004	SPI2	2	A	13-JAN-2004	9	26-JAN-2004	9	1931	0	M200

jahr	betrieb	seq_u	statf	eintritt	ein_s	austritt	aus_s	born_y	doppel	v14_01
2002	SPI3	3	A	15-FEB-2002	17	25-FEB-2002	17	1964	0	M100
2002	SPI3	.	A	06-MAY-2002	15	21-MAY-2002	11	1965	5	M300
2002	SPI3	4	A	06-MAY-2002	16	21-MAY-2002	11	1965	6	M300
2002	SPI3	5	A	11-SEP-2002	14	15-SEP-2002	16	1964	0	M700

In weiteren Fällen ist eine Mehrfachmeldung möglich, aber aus den Daten nicht eindeutig zu erkennen (vgl. Kap. 4.2).

Aufgeschlüsselt nach Kalenderjahren zeigt sich, dass Mehrfachmeldungen seit 2009 deutlich seltener sind als in der Zeit davor (Tab. 8), mit einer weiteren Verbesserung für die Datenjahrgänge ab 2014.

Tab. 8 Potenzielle Mehrfachmeldungen von Spitalaufenthalt, nach Übereinstimmungskriterium

	"sichere" Mehrfachmeldung (Tab.7b)	selbe Eintritts- stunde, selbes Austrittsdatum (ohne Std.)	"wahrscheinliche" Mehrfachmeldung	selbes Eintrittsdatum (ohne Std.), selbe Austrittsstunde	"mögliche" Mehrfachmeldung	"eher keine" Mehrfach- meldung	potenzielle Mehrfachmel- dungen total	keine Mehrfach- meldung	Total
Jahr									
2000	2'445	745		292	2'746	393	6'621	1'177'608	1'184'229
2001	1'477	1'245		220	1'945	405	5'292	1'297'672	1'302'964
2002	4'616	1'205		297	1'556	361	8'035	1'338'094	1'346'129
2003	10'054	412		205	1'412	353	12'436	1'378'467	1'390'903
2004	831	358		222	1'294	289	2'994	1'406'749	1'409'743
2005	705	103		171	885	825	2'689	1'463'778	1'466'467
2006	2'757	85		249	1'081	803	4'975	1'495'689	1'500'664
2007	560	79		199	475	933	2'246	1'536'077	1'538'323
2008	2'233	73		98	451	580	3'435	1'581'319	1'584'754
2009	479	40		52	219	145	935	1'327'192	1'328'127
2010	272	57		70	271	121	791	1'344'454	1'345'245
2011	246	88		39	445	132	950	1'362'748	1'363'698
2012	141	2		16	784	175	1'118	1'352'403	1'353'521
2013	416	12		8	439	88	963	1'373'491	1'374'454
2014	113	10		16	261	74	474	1'400'540	1'401'014
2015	54	2		12	250	84	402	1'429'799	1'430'201
2016	129	10		12	169	110	430	1'467'517	1'467'947
Total	27'528	4'526		2'178	14'683	5'871	54'786	23'733'597	23'788'383

3 Die MedStat wieder konsistent machen

Die Identifizierung von fehlerhaften AVCs, B-/C-Fällen ohne Folgehospitalisierung, zeitlich überlappenden Meldungen und Mehrfachmeldungen² hat Folgen für die Konsistenz der gesamten Datenbasis:

1. Die Sequenznummern sind nicht mehr immer korrekt.
2. Die Spitalaufenthaltsdauer ist möglicherweise nicht mehr korrekt.
3. Die Zeit zwischen den Spitalaufenthalten ist möglicherweise nicht mehr korrekt.

Für die Bereinigung wurden folgende neuen Variablen definiert:

DAUER_U* Aufenthaltsdauer (in Tagen): Austrittsdatum-Eintrittsdatum

Wertebereich: 0 - 32'990

SEQ_U Sequenznummer MedStat** (Reihenfolge nach Eintrittsdatum, -stunde)

alle Meldungen werden gezählt ausser solche von ambulanten und teilstationären Aufenthalten, von sicheren oder als wahrscheinlich eingeschätzten Mehrfachmeldungen, von B-/C-Fällen mit Folgehospitalisierung und solche mit Missing oder als ungültig erachteten AVCs (incl. AVC = 0)

Wertebereich: 1 - 334

SEITLASTHOSP Tage seit letztem Spitalaustritt

Wertebereich: -19'082 - 6197

OVERLAP* Hospitalisierung ist überlappend** mit einer vorhergehenden

0 (keine Überlappung)

1 eingeschlossen von anderer Hospitalisierung

2 teilweise (Beginn ist) überlappend mit anderer Hospitalisierung

OVERLAPREST** Anzahl Tage, die nicht von der vorhergehenden Hospitalisierung abgedeckt sind**

Wertebereich: 0 - 32'990

SAMESPIT gleiches oder anderes Spital

(bezogen auf die letzte vorhergehende Hospitalisierung)

0 keine frühere Hospitalisierung oder anderes Spital

1 war gleiches Spital

LASTHOSP Hospitalisierung mit spätestem Austrittsdatum

(14'632'301 missings: keine weiteren gültigen Hospitalisierungen unter diesem AVC)

0 (ist nicht letzte Hospitalisierung dieses AVCs)

1 Hospitalisierung hat das späteste Austrittsdatum dieses AVCs

2 B- oder C-Fall, der gleichzeitig das späteste Austrittsdatum dieses AVCs aufweist (notwendigerweise ein B-/C-Fall ohne Folgehospitalisierung)

² Da nicht immer eindeutig feststellbar ist, ob zwei oder mehr Hospitalisierungen mit identischem Datum Mehrfachmeldungen desselben Aufenthaltes sind oder mindestens eine Meldung ein falsches Eintritts- und/oder Austrittsdatum aufweist, wurden die nachstehend erwähnten Variablen (Sequenznummer, Tage seit letztem Spitalaustritt, "Hospitalisierung ist überlappend", "nicht von einer vorhergehenden Hospitalisierung abgedeckte Tage" und "gleiches oder anderes Spital wie bei vorhergehender Hospitalisierung") nur bei den "sicheren" und den "wahrscheinlichen" Mehrfachmeldungen angepasst. Bei "möglicher Mehrfachmeldung" und "eher keine Mehrfachmeldung" kann eine Entscheidung erst aufgrund der konkreten Fragestellung/Auswertungsstrategie und den – uns in dieser Phase nicht zur Verfügung stehenden – Diagnosen und Behandlungen getroffen werden.

- * DAUER_U: Die MedStat selbst liefert zwei Variablen für die Spitalaufenthaltsdauer, die Dauer nach Swiss-DRG und die Dauer gemäss OECD-Definition, wobei die Swiss-DRG-Dauer in den meisten Fällen einen Tag kürzer ist als die der OECD. Da wir die Dauer in Kalendertagen und nicht eine anderswo definierte Dauer benötigen, haben wir die Differenz zwischen Aus- und Eintrittsdatum berechnet.
- ** SEQ_U: Eine Sequenznummer ist in der MedStat zwar vorhanden, sie fehlt aber bei B-/C-Fällen, da diese einen Fortsetzungsrecord im Folgejahr haben müssten. Da dieser Fortsetzungsrecord in vielen Fällen aber fehlt und dann die Hospitalisierung nicht gezählt würde (und weil wir auch Doppelmeldungen und Meldungen mit fehlerhaften oder fehlendem AVCs ausschliessen mussten), haben wir eine neue Sequenznummer generiert, die auch B-/C-Hospitalisierungen ohne Fortsetzungsrecord zählt, dafür Mehrfachmeldungen und Meldungen mit fehlerhaftem AVC ausschliesst. Um geografisch und zeitlich konsistente Daten zu erhalten, wurden ambulante und teilstationäre Aufenthalte bei der Sequenznummer nicht mitgezählt.
- *** OVERLAP: Gelegentlich kommt es vor, dass Spitalaufenthalte überlappend oder deckungsgleich sind (aus administrativen Gründen oder wegen Datenfehlern): Dies wird mit dieser Variable festgehalten.
- **** OVERLAPREST: Um in diesen Fällen die Gesamtaufenthaltsdauer nicht doppelt zu zählen, wird für jeden Spitalaufenthalt festgehalten, wie viele Tage von keinem Aufenthalt mit früherem Eintrittsdatum überlappt werden.

4 Konstellationen, die mit "vernünftigem" Aufwand nicht lösbar sind

4.1 B-/C-Fälle ohne Fortsetzungsrecord

Bei B-/C-Fällen ohne Fortsetzungsrecord im Folgejahr sollte geprüft werden, ob

- (A) es wirklich(e) B-/C-Fälle sind. Falls ja, was mit dem Fortsetzungsrecord geschehen ist (denkbar sind z.B. bei Spitalfusionen Fortsetzungsrecords mit einer neuen Betriebsnummer aber Eintrittsdatum 1. Januar des Folgejahres).
- (B) es sich in Wirklichkeit um einen normaler A-Fall mit Austritt handelt. Es gibt Hinweise, dass nicht alle B-Fälle wirklich B-Fälle sind, sondern vielmehr die Austrittsmeldung "untergegangen" ist womit die Person dann Ende Jahr automatisch als B-Fall registriert wurde (Ein solcher Hinweis sind z.B. AVCs mit zwei Hospitalisierungen am selben Tag, wobei eine ein A-Fall ist, die andere hingegen ein B-Fall) . Hier könnte es sich um ein Duplikat einer Meldung handeln (der korrekte Austritt wurde gemeldet, das Duplikat (der B-Fall) und dessen fehlender Fortsetzungsrecord deutet bei diesem B-Fall auf eine "Karteileiche". Dieser Eindruck verstärkt sich, wenn auch die Eintrittsstunde übereinstimmt, oder – wie im folgenden Beispiel – noch eine weitere, spätere Hospitalisierung im selben Jahr hinzukommt).

Jahr	Betrieb	Staff	Eintritt /-stunde		Austritt /-stunde		Born	Hauptkostenstelle
2001	SPI1	A	11. Apr 2001	7	13. Apr 2001	9	1947	M200
2001	SPI2	A	20. Juni 2001	10	2. Juli 2001	0	1947	M200
2001	SPI2	B	20. Juni 2001	10	bis Ende Jahr im Spital		1947	M200
2001	SPI2	A	17. Juli 2001	8	19. Juli 2001	0	1947	M200
2002	SPI2	A	17. Juli 2002	8	24. Juli 2002	14	1947	M200
2003	SPI3	A	5. Feb 2003	10	19. Feb 2003	10	1947	M200

- (C) Im neuen oder im alten Jahr ein fehlerhafter AVC erstellt wurde.

Da wir nicht bei allen B-/C-Fällen ohne Fortsetzungsrecord über solche Informationen verfügen, können wir nicht immer herausfinden, was wirklich korrekt ist. Zu den 80'462 B-/C-Fällen ohne Fortsetzungsrecord (Tab. 6) gibt es nur 24'466 Meldungen mit mindestens einem späteren Eintrittsdatum desselben Patienten. Es gibt 3'873 B-/C-Fälle ohne Fortsetzungsrecord, die aber ein Eintrittsdatum haben, das früher ist als das Austrittsdatum der vorhergehenden Hospitalisierung (Verdacht: der Patient wurde entlassen ohne dass dies entsprechend registriert wurde, womit ein B-Fall als Karteileiche entstand). 7'978 Meldungen betreffen Spitaleintritte, deren Eintrittsdatum im selben Jahr aber später als das Eintrittsdatum der B-/C-Meldung ist (Verdacht: Der Patient wurde entlassen und ist später im Jahr wieder hospitalisiert worden). Dies in jedem Einzelfall abzuklären, ist mit vernünftigem Aufwand nicht möglich.

4.2 "Mögliche" Mehrfachmeldungen

Im Gegensatz zu den "sicheren" bzw. "wahrscheinlichen" Mehrfachmeldungen (vgl. Kap. 2.3) ergeben sich bei den **"möglichen"** Mehrfachmeldungen oft kaum lösbare Situationen.

Beispiel 1: Übereinstimmendes Eintritts- und Austrittsdatum, aber sowohl Eintritts- als auch Austrittsstunde jeweils verschieden. Das könnte ein Duplikat sein, es wäre aber auch denkbar, dass a) die behandelte Person am 26. April um 15h auf eine andere Abteilung verlegt und am 27. April um 9 Uhr wieder zurückverlegt wurde, was eine vollständige Überlappung von zwei Spitalaufenthalten bedeuten würde; allerdings bleibt es ein Rätsel, dass die Hauptkostenstelle dieselbe ist. Denkbar ist auch ein Datumsfehler, d.h., die Person ist am 26. April um 14h ausgetreten, und dann um 15h wieder eingetreten.

jahr	betrieb	seq_u	statf	eintritt	ein_s	austritt	aus_s	born_y	doppel	v14_01
2010	SPI1	1	A	30-MAR-2010	12	02-APR-2010	11	1950	0	M200
2010	SPI2	2	A	26-APR-2010	8	27-APR-2010	14	1950	10	M200
2010	SPI2	3	A	26-APR-2010	15	27-APR-2010	9	1950	9	M200
2010	SPI1	4	A	21-JUN-2010	7	23-JUN-2010	10	1949	0	M200

Beispiel 2: Nahtlose Aneinanderreihung von 4 Hospitalisierungen seit dem 20. November 2011 mit B-Fall als Abschluss. Bei der Fortsetzungshospitalisierung im Folgejahr ist aber nicht das Eintrittsdatum des B-Falles angegeben, auch nicht dasjenige der ersten der 4 nahtlosen Hospitalisierungen, sondern dasjenige der zweiten. Unsere Interpretation: 1. Hospitalisierung unter M100 (Innere Medizin), gefolgt von M900 (Geriatric-)Aufenthalt; mit Kurzverlegung in M000 (Fachgebiete allgemein) – da der Record mit der Sequenznummer (seq_u) 29 vermutlich auch die Behandlungen und Diagnosen von seq_u 28 enthält, wäre man mit einer Löschung der seq_us 28 und 31 am besten bedient.

Was in diesem Fall nur "glücklicherweise" so ist: seq_u 28 wird zwar wegen der Identifizierung als Doppelmeldung "gelöscht", seq_u 31 jedoch zunächst nicht, da es zu dieser B-Meldung keinen korrekten Fortsetzungsrecord gibt; seq_us 28, 30 und 31 werden allerdings erkannt als Hospitalisierungen, die vollständig von einer anderen Hospitalisierung (seq_u 29 – dem eigentlichen Fortsetzungsrecord von seq_u 4) umschlossen sind.

jahr	betrieb	seq_u	statf	eintritt	ein_s	austritt	aus_s	born_y	doppel	v14_01
2011	SPI3	27	A	20-NOV-2011	23	12-DEC-2011	13	1939	0	M100
2011	SPI3	28	A	12-DEC-2011	13	19-DEC-2011	15	1939	10	M900
2011	SPI3	30	A	19-DEC-2011	15	20-DEC-2011	0	1939	0	M000
2011	SPI3	31	B	20-DEC-2011	0	.	.	1939	0	M900
2012	SPI3	29	A	12-DEC-2011	13	07-JAN-2012	2	1939	9	M900

Beispiel 3: ähnliche Situation wie Beispiel 2: Ein Patient mit 7 Aufenthaltsmeldungen zwischen 17. Juni 2011 und 14. Juni 2012, zum Teil nahtlos aneinandergereiht (inklusive Aus- und Eintrittsstunden); seq_us 17, 19 (Geriatric) und 20 (Chirurgie) sind vollständig überlappt durch seq_u 21 (Rehabilitation) – eigentlich der Fortsetzungsrecord von B-Fall seq_u 19, aber wegen früheren Eintrittsdatums von A-Fall seq_u 17 nicht als Fortsetzungsrecord erkannt wurde.

Der Patient war jedenfalls ununterbrochen im gleichen Spital in verschiedenen Abteilungen; weshalb jeweils Ein- und Austritte gemeldet wurden, bleibt spekulativ. Völlig unklar ist, warum die Geriatric am 6. Oktober einen Austritt meldete und zur gleichen Stunde wieder einen Eintritt, welcher dann zu einem B-Fall wurde, aber ohne Fortsetzung auf der Geriatric. Ebenfalls rätselhaft ist, weshalb beim ersten Austritt 2012 (seq_u 18) das Eintrittsdatum auf den 19. Juli 2011 gesetzt wurde; es könnte vermutet werden, dass die Reha wusste, dass die

Person direkt aus der Geriatrie gekommen ist, aber nicht, dass der Geriatrieaufenthalt unmittelbar einem Aufenthalt auf der Chirurgie gefolgt war.

jahr	betrieb	seq_u	statf	eintritt	ein_s	austritt	aus_s	born_y	doppel	v14_01
2011	SPI4	14	A	17-JUN-2011	13	18-JUN-2011	14	1938	0	M000
2011	SPI4	15	A	18-JUN-2011	14	12-JUL-2011	11	1938	0	M950
2011	SPI4	16	A	12-JUL-2011	11	19-JUL-2011	9	1938	0	M200
2011	SPI4	17	A	19-JUL-2011	9	06-OCT-2011	10	1938	10	M900
2011	SPI4	19	B	06-OCT-2011	10	.	.	1938	0	M900
2012	SPI4	20	A	13-FEB-2012	11	16-FEB-2012	11	1939	0	M200
2012	SPI4	18	A	19-JUL-2011	9	14-JUN-2012	14	1938	9	M950
2012	SPI4	21	A	26-SEP-2012	7	29-SEP-2012	9	1938	0	M200
2012	SPI4	.	B	29-SEP-2012	9	.	.	1938	0	M950
2013	SPI4	22	A	29-SEP-2012	9	24-JAN-2013	13	1938	0	M900
2015	SPI4	23	A	05-OCT-2015	13	07-OCT-2015	15	1938	0	M100

Beispiel 4: Zwei Hospitalisierungen am 29. Dezember morgens um 4 Uhr mit Austrittsdatum 31. Dezember 13 Uhr – "eine" Geburt: Korrekt vom Algorithmus als "sichere" Doppelmeldung erkannt. Dann zwei weitere Hospitalisierungen im Folgejahr in einem anderen Spital (oder demselben mit anderer Spitalnummer, z. B. wegen Spitalfusion?), mit Eintrittsdatum 29. Dezember aber Eintrittsstunde 11h bzw. 20h und Austrittsdatum 4. Januar 10h. Ebenfalls vom Algorithmus als "wahrscheinliche Doppelmeldung" erkannt.

Gemäss unserer primären Einschätzung handelte es sich um eine einzige Hospitalisierung (Geburt und dann wegen Komplikationen auf eine andere Abteilung verlegt), die korrekterweise als B-Fall im Jahr 2003 und Folgefall im Jahr 2004 hätte gemeldet werden müssen (es existiert aber kein B-Fall zu diesem AVC). Berücksichtigt man jedoch die Informationen der Neugeborenen-Zusatzdaten, zeigt sich, dass es sich um eine Doppelmeldung handelt (seq_u 1 und 2), plus zwei weitere Geburten mit jeweils falschem AVC.

betrieb	seq_u	eintritt		austritt		doppel	Haupt- kosten- stelle	Neugeborenen-Zusatzdaten:					
								Geb. zeit	Vital- status	Mehr- ling	Rang	Ge- wicht	Grös- se
SPI5	1	29-DEC-2003	4	31-DEC-2003	13	2	M300	04:25	1	1	1	3090	50
SPI5	2	29-DEC-2003	4	31-DEC-2003	13	1	M300	04:25	1	1	1	3090	50
SPI6	3	29-DEC-2003	11	04-JAN-2004	10	6	M300	11:19	1	1	1	3090	53
SPI6	4	29-DEC-2003	20	04-JAN-2004	10	5	M300	20:47	1	1	1	3560	48

Fazit:

Algorithmen zu entwickeln, die solche Konstellationen korrekt erkennen und nicht fälschlicherweise ähnliche, aber in Realität nicht zusammengehörige Records verbinden, ist sehr aufwändig. Die Alternative, alle ähnlichen Situationen im Einzelnen durchzusehen und zu entscheiden, ist wegen des grossen Zeitaufwands unrealistisch, denn es handelt sich dabei nicht um ein paar Dutzend, sondern um mehrere tausend Fälle.

4.3 Komplexe Spitalaufenthalt-Serien

Grundsätzlich sind verschachtelte Spitalaufenthalte schwierig aufzulösen (also z.B. Langzeitaufenthalte in Psychiatrie / Geriatrie mit zwischenzeitlichen Hospitalisierungen in Chirurgie ohne Austrittsmeldung). Wie das folgende komplexe Beispiel (Tab. 9) zeigt, ist es nicht immer eindeutig nachzuvollziehen, wie diese Spitalaufenthalte in eine korrekte Beziehung zueinander gesetzt werden können.

Tab. 9 Exemplarische Darstellung einer komplexen Patienten-Geschichte

(aus Zellweger&Bopp 2017; Alter und Geschlecht des Patienten / der Patientin stimmen bei allen Aufenthalten überein)

Nr	Stat- fall	Jahr	Betrieb	Eintritt	Austritt	Bemerkungen
1	A	1998	z	31-AUG-1998	31-AUG-1998	
2	C	1999	yy	26-JUL-1993	.	redundant wegen 3
3	C	2000	yy	26-JUL-1993	.	A)
4	B	2003	yy	01-JAN-2003	.	redundant wegen 5
5	C	2004	yy	01-JAN-2003	.	redundant wegen 10
6	A	2004	v	06-MAR-2004	07-MAR-2004	
7	A	2004	p	12-AUG-2004	13-AUG-2004	
8	C	2005	yy	01-JAN-2003	.	redundant wegen 10
9	A	2006	w	19-JAN-2006	25-JAN-2006	
10	A	2006	yy	01-JAN-2003	31-DEC-2006	B)
11	B	2007	yy	01-JAN-2007	.	redundant wegen 12
12	A	2008	yy	01-JAN-2007	31-DEC-2008	C)
13	A	2010	u	17-JUL-2010	18-JUL-2010	
14	A	2010	r	16-OCT-2010	17-OCT-2010	
15	A	2011	yy	01-JAN-2009	03-JAN-2011	
16	A	2011	yy	04-MAR-2011	03-OCT-2011	
17	A	2011	w	21-APR-2011	24-APR-2011	
18	A	2011	v	09-OCT-2011	11-OCT-2011	
19	B	2011	yy	16-DEC-2011	.	redundant wegen 21
20		2011	-	19-DEC-2011	20-DEC-2011	Heimaufenthalt
21	A	2012	yy	16-DEC-2011	13-FEB-2012	
22	A	2012	v	30-JAN-2012	01-FEB-2012	
23	A	2012	x	01-FEB-2012	03-FEB-2012	
24	A	2012	x	05-MAR-2012	12-MAR-2012	
25	A	2012	yy	14-FEB-2012	07-AUG-2012	
26	A	2012	t	11-JUN-2012	11-JUN-2012	
27	A	2012	x	11-JUN-2012	15-JUN-2012	
28	B	2012	S	07-AUG-2012	.	
29	A	2012	x	04-DEC-2012	13-DEC-2012	
30	A	2012	x	24-DEC-2012	31-DEC-2012	D)

Es lassen sich folgende Rückschlüsse zu den vier Fällen A) bis D) ziehen:

- A)** Dieser Fall müsste eigentlich ebenfalls redundant sein. Es gibt aber keinen Fortsetzungsrecord im Jahr 2001. Verbleibt deshalb im Datensatz. Die Frage bleibt: Wo war diese Person 2001 und 2002?
- B)** Ist mit aller Wahrscheinlichkeit kein A-Fall, sondern ein B-Fall (wahrscheinlich sogar C-Fall), könnte wie Nr. 8 redundant sein und mit Nr. 15 einen Fortsetzungsrecord haben, dann aber müsste bei Nr. 15 das Eintrittsdatum auf 1. Januar 2003 gesetzt werden, was aber aller Wahrscheinlichkeit nach gleich selbst ein Fortsetzungsrecord einer unbekannten Hospitalisierung des Vorjahres 2002 ist (siehe Bemerkung A).
- C)** Ist mit aller Wahrscheinlichkeit kein A-Fall, sondern ein (Fortsetzungs-) C-Fall von Nr. 11, von dem es selbst wieder eine Fortsetzung gibt (Nr. 15) und der somit redundant wäre.
- D)** Ist vielleicht ebenfalls ein B-Fall und kein Austritt – bei dieser Vorgeschichte! Allerdings gibt es in diesem Jahr für diese Person bereits einen B-Fall eines anderen Spitals (Nr. 28). Was in der Realität zutrifft, lässt sich nicht eindeutig entscheiden.

5 Inkonsistenzen in der SOMED

5.1 Fehlerhafter Anonymer Verbindungscode (AVC)

Die SOMED weist für die Jahre 2007 bis 2016 1'955'913 Meldungen mit insgesamt 578'861 verschiedenen AVCs auf. Bei 390 Meldungen fehlt der AVC; zwei Verbindungscode im Jahr 2007 scheinen ein "Default-Code" zu sein, einer für Männer, der 707 mal vorkommt, bei dem das Geburtsjahr und das Geschlecht (mit einer Ausnahme) immer übereinstimmen, der aber Einwohner von 12 Kantonen betrifft von 11 Betrieben gemeldet wurde und 128 Mal als Austrittsgrund "Tod" angibt; einer für Frauen (72 Meldungen). Diese Codes werden im Folgenden als ungültig erachtet (keine Sequenznummer vergeben, Spezialcode 9 in Variable FOLGEHEIM).

Anders als in der MedStat ist in der SOMED der AVC nur selten von Inkonsistenzen betroffen, obwohl, in der SOMED jedes Jahr für jeden im Laufe des Jahres betreuten Klienten eine Meldung gemacht werden muss.

Zieht man den oben erwähnten "Default"-Code ab, dann sind nur 564 AVCs mit insgesamt 3'065 Meldungen von Inkonsistenzen (Unterschiede im Geschlecht und oder Geburtsjahr) betroffen (Tab. 10). Bei keiner dieser Meldungen ist jedoch die Differenz der Geburtsjahre grösser als 1 (im Gegensatz zur MedStat ist in der SOMED eine Abweichung von einem Jahr eine Inkonsistenz, da die Variable Alter hier als – im Falle von Austritten oder Todesfällen auch fiktives – Alter am Ende des Kalenderjahres definiert ist) und bei nur 5 Meldungen ist das Geschlecht nicht identisch. Wir gehen daher davon aus, dass der AVC (ausser in den beiden oben genannten Fällen) in der SOMED immer korrekt ist.

Tab. 10 Anzahl AVCs in der SOMED, nach Jahr

jahr	Total	kein AVC	primär ungültiger AVC	primär gültiger AVC	Meldungen mit potenziell inkonsistentem AVC
2007	183'631	309	779	182'543	621
2008	186'328	81	0	186'247	541
2009	190'352	0	0	190'352	426
2010	189'117	0	0	189'117	329
2011	190'692	0	0	190'692	278
2012	195'931	0	0	195'931	237
2013	200'875	0	0	200'875	194
2014	204'445	0	0	204'445	163
2015	211'646	0	0	211'646	148
2016	202'896	0	0	202'896	128
Total	1'955'913	390	779	1'954'744	3'065

5.2 Fehlende Fortsetzungsrecords im Folgejahr

Da für die SOMED jedes Jahr alle Klienten gemeldet werden müssen, die im Laufe des Jahres im Betrieb registriert waren, gibt es für Klienten, die über das Jahresende hinaus betreut wurden mindestens zwei Meldungen, bei langjährigen Klienten sogar mehrere.

Zuerst wurde geprüft, ob es für die Klienten ohne Austrittsmeldung einen Fortsetzungsrecord gibt:

- im ersten Schritt war das Kriterium dafür dasselbe Eintrittsdatum im selben Heim;

- im zweiten Schritt wurde ein anderer Betrieb, nicht aber ein abweichendes Eintrittsdatum akzeptiert;
- im dritten Schritt wurde eine Eintrittsdifferenz von maximal 31 Tagen akzeptiert (da Stichproben gezeigt hatten, dass manchmal des Eintrittsdatum in späteren Jahren leicht variierte).

Schliesslich wurde festgehalten ob es für Aufenthalte ohne Austrittsmeldung in den Folgejahren irgendeinen späteren SOMED-Aufenthalt gibt oder ob es das letzte gemeldete Datum ist (Tab. 11a & 11b). Dies gilt für die Jahre 2007 bis 2015 – naturgemäss ist es für das Jahr 2016 noch nicht möglich zu wissen, ob es im Jahr 2017 einen Fortsetzungsrecord gibt.

Tab. 11a Heimaufenthalte mit und ohne Austrittsdatum

	AVC=0 oder Default-AVC" von 2007	mit Austrittsdatum	ohne Austrittsdatum			Total
			Mit Fortsetzungsrecord	Ohne Fortsetzungsrecord	Subtotal	
2007	1'088	57'461	120'785	4'297	125'082	183'631
2008	81	58'707	123'495	4'045	127'540	186'328
2009	0	59'941	126'100	4'311	130'411	190'352
2010	0	58'520	125'964	4'633	130'597	189'117
2011	0	57'962	127'519	5'211	132'730	190'692
2012	0	61'234	130'538	4'159	134'697	195'931
2013	0	63'735	132'730	4'410	137'140	200'875
2014	0	65'387	133'457	5'601	139'058	204'445
2015	0	70'440	129'245	11'961	141'206	211'646
2016	0	68'430	*	*	134'466	202'896
Total	1'169	621'817	1'149'833	48'628	1'332'927	1'955'913

* erst aus den Daten des uns nicht vorliegenden Jahrgangs 2017 bestimmbar

Tab. 11b Aufenthalte ohne Austrittsdaten (am Jahresende noch im Heim)

	mit Fortsetzungsrecord im Folgejahr				ohne Fortsetzungsrecord im Folgejahr			
	übereinstimmendes Eintrittsdatum, selber Betrieb	übereinstimmendes Eintrittsdatum, aber anderer Betrieb	Eintrittsdifferenz von max. 30 Tagen zugelassen	Subtotal	zugleich letzte SOMED-Meldung	Eintritt in einem Vorjahr, taucht später wieder auf	Eintritt im Meldungsjahr, taucht später wieder auf	Subtotal
2007	119'854	620	311	120'785	2'845	859	593	4'297
2008	122'294	909	292	123'495	2'497	981	567	4'045
2009	124'665	1'089	346	126'100	2'743	925	643	4'311
2010	123'951	1'537	476	125'964	2'873	1120	640	4'633
2011	125'994	985	540	127'519	2'152	2406	653	5'211
2012	128'908	1'050	580	130'538	2'215	1287	657	4'159
2013	130'889	1'316	525	132'730	2'365	1291	754	4'410
2014	132'123	788	546	133'457	2'552	2253	796	5'601
2015	127'899	801	545	129'245	9'371	1784	806	11'961
Total	1'136'577	9'095	4'161	1'149'833	29'613	12906	6'109	48'628

Auffällig sind die 9'371 Aufenthalte im Jahr 2015 ohne irgendeinen weiteren Folgeaufenthalt (rund vier mal so viele Meldungen als in allen Vorjahren). Es fällt auf, dass sie gehäuft in Betrieben auftreten, die 2016 offenbar nicht mehr existierten. Da es für alle diese Klienten keine 2016er Meldung gibt, muss offen bleiben, was hier geschehen ist. Der Verdacht, es könnten zu viele Meldungen sein, liegt nahe, weil 2015 nicht nur rund 7'000 Meldungen mehr gemacht wurden als in den Vorjahren sondern auch als im Folgejahr 2016 (siehe letzte Spalte Tab. 12a). Eine exploratorische Analyse zeigt, dass für fast alle diese Personen auch schon 2014 eine Meldung ohne Austrittsdatum gemacht worden war. Erklärte man diese Records für ungültig, würde das Problem deshalb nicht gelöst, sondern nur um ein Jahr vorverschoben.

5.3 Neue Variablen aus der Bereinigung von MedStat und SOMED

Folgende zusätzliche Variablen wurden gebildet:

DAUER_U Aufenthaltsdauer (in Tagen): Austrittsdatum-Eintrittsdatum Wertebereich: 0 - 54'292
SEQ_U: Sequenznummer (Reihenfolge nach Eintrittsdatum)* alle Meldungen werden gezählt, ausser Fälle ohne Austritt, aber mit Folgeaufenthalt(en) Wertebereich: 1 - 127
SEITLASTHEIM Tage seit letztem Heimaufenthalt: Eintrittsdatum-Austrittsdatum des vorherigen Aufenthaltes Wertebereich: -27'981 - 3'628
OVERLAP: Aufenthalt ist überlappend mit einem vorhergehenden** 0 (keine Überlappung) 1 eingeschlossen von 2 teilweise (Beginn ist) überlappend
OVERLAPREST: Anzahl Tage, die nicht vom vorhergehenden Heimaufenthalt abgedeckt sind*** Wertebereich: 0 - 54'292
LASTHEIM: Letzter Heimaufenthalt 0 (ist nicht spätester Aufenthalt) 1 Aufenthalt hat das späteste Heimaustrittsdatum dieses AVCs
FOLGEHEIM Fortsetzungsaufenthalt im Folgejahr -1 hat Austrittsdatum oder ist ein 2016er Record 0 kein Fortsetzungsrec, Eintritt im Vorjahr, taucht später wieder auf 1 kein Fortsetzungsrec, Eintritt im Meldejahr, taucht später wieder auf 2 kein Fortsetzungsrec, zugleich spätester SOMED-Eintritt 3 hat Fortsetzungsrec (max. 30 Tage später, gleiches Heim) 4 hat Fortsetzungsrec im gleichen Heim 5 hat Fortsetzungsrec in anderem Heim 9 AVC=0 oder Default-AVC" von 2007.

- * SEQ_U: Eine Sequenznummer wird in der SOMED nicht mitgeliefert. Da jedes Jahr für jeden Klienten ein Record generiert wird, können die Records, welche kein Austrittsdatum haben und im Folgejahr ununterbrochen wiedergemeldet werden, nicht als eigener Heimaufenthalt, sondern als Fortsetzung des Heimaufenthaltes betrachtet werden.
- ** OVERLAP: Gelegentlich kommt es vor dass Heimaufenthalte überlappend oder deckungsgleich sind (aus administrativen Gründen oder wegen Datenfehlern): Diese ist hier festgehalten.
- *** OVERLAPREST: Um in diesen Fällen die Gesamtaufenthaltsdauer nicht doppelt zu zählen wird für jeden Heimaufenthalt festgehalten, wieviele Tage von einem früheren Heimaufenthalt NICHT überlappt sind.

6 MedStat und SOMED kombiniert

6.1 Kennzeichnung überlappender Heim-/Spitalaufenthalte

Spital- und Heimaufenthalte können nicht nur innerhalb des gleichen Institutionstyps überlappen, sondern auch untereinander. Spital- und Heimaufenthalte dürfen daher nicht einfach zusammengezählt werden sondern müssen so aufbereitet werden, dass Doppelzählungen von Aufenthaltstagen ausgeschlossen werden können und eine eindeutige Reihenfolge bestimmt werden kann.

Dafür werden weitere Variablen benötigt.

Einerseits interessiert aus Gründen des Follow-up die Zeitdauer bis zum letzten dokumentierten Spitalaufenthalt (TAGE_BIS_LAST_HOSP), andererseits erwiesen sich die Variablen von MedStat und SOMED zum Aufenthaltsort vor und nach dem Spital- bzw. Heimaufenthalt als relativ unzuverlässig. Direkte Übergänge von und zu Spitälern bzw. Heimen wurden in den zwei Variablen: VONORG, ZUORG, festgehalten, wobei bei den Heimeintritten aus einem Spital bzw. den Austritten aus einem Spital in ein Heim eine Karenzfrist von 10 Tagen zugelassen wurde.

Da bei Heimaufenthalten Verlegungen ins Spital zur Behandlung akuter Erkrankungen mit geplanter Rückkehr ins Heim nicht als Heimaustritte behandelt werden (sofern das Bett/das Zimmer vom Klienten belegt bleibt), muss für jeden Heimaufenthalt dokumentiert sein, an wie vielen Tagen dieses Aufenthaltes der Klient auch in einem Spital gemeldet war (SPIOVER). Damit wird verhindert, dass die Aufenthaltsdauer in Institutionen doppelt gezählt wird. Dies ist korrekt, insoweit der einzelne Heimaufenthalt betrachtet wird; überlappen sich allerdings Heimaufenthalte, welche wiederum von einem Spitalaufenthalt überdeckt werden können die Tage nicht bei beiden Heimaufenthalten abgezählt werden, sondern nur bei einem (SPIOVER2).

Beispiel: Eine Person mit sechs Aufenthalten in drei verschiedenen Heimen (HEI1, HEI2, HEI3) und mit drei Aufenthalten in zwei Spitälern (SPI1, SPI2):

Heimaufenthalt 1 überlappt vollständig die Heimaufenthalte 2, 3 und 4 (darum: OVERLAPREST=0, also keine anrechenbare Aufenthaltstage). Heimaufenthalt 5 ist nur im Jahr 2009 nicht durch Heimaufenthalt 1 überlappt und es verbleiben 365 anrechenbare Aufenthaltstage (OVERLAPREST). Sowohl Heimaufenthalt 1 als auch 5 überlappen mit zwei Spitalaufenthalten 6 und 7 um insgesamt 21 Tage.

Korrekterweise werden bei SPIOVER bei beiden Heimaufenthalten, die von der Überlappung betroffen sind, die 21 Tage vermerkt. Betrachtet man hingegen den gesamten Aufenthalt in stationären Einrichtungen (OVERLAP2), dürfen die 21 Tage bei Heimaufenthalt 5 nicht mitgezählt werden, da diese bereits vom Heimaufenthalt 1 abgezogen worden sind.

(Bei Heimaufenthalt 8, der ebenfalls noch um über ein Jahr mit Heimaufenthalt 5 überlappt (ein Teil 2008 und ganz 2009), kann korrekt die Dauer von Spitalaufenthalt 9 – ein Tag – abgezogen werden).

Nr	jahr	betrieb	dauer	overlaprest	spiover	spiover2	eintritt	austritt
1	2008	HEI1	1032	1032	21	21	06-MAR-2006	kein Austritt
2	2007	HEI2	198	0	0	0	04-APR-2007	19-OCT-2007
3	2008	HEI2	10	0	0	0	20-MAR-2008	30-MAR-2008
4	2008	HEI3	39	0	0	0	15-SEP-2008	24-OCT-2008
5	2009	HEI2	652	365	21	0	20-MAR-2008	kein Austritt
6	2008	SPI1	2	2	0	0	25-AUG-2008	27-AUG-2008
7	2008	SPI2	19	19	0	0	27-AUG-2008	15-SEP-2008
8	2012	HEI2	1237	804	1	1	24-OCT-2008	14-MAR-2012
9	2012	SPI1	1	1	0	0	11-MAR-2012	12-MAR-2012

6.2 Neue Variablen aus der Kombination von MedStat und SOMED

INVALID: *Zusammenfassung der Gründe für Ungültigkeitserklärung (bei mehreren Gründen für eine Ungültigerklärung überschreibt die höhere Zahl die niedrigere)*

- 0 Record ist gültig/valide
- 2 AVC=0
- 3 Späterer Heimrecord mit gleichem Eintrittsdatum vorhanden
- 4 sichere/wahrscheinliche Doppelmeldung
- 5 B-/C-Fall mit nachfolgendem A-Fall
- 6 ambulante oder teilstationäre Hospitalisierung
- 7 ungültiger AVC.

TAGE_BIS_LAST_HOSP: *Tage bis zu letztem Spitaleintritt* (letzter Spitaleintritt minus Austritt des aktuellen Aufenthaltes – wenn negativ, dann erfolgte der letzte Spitaleintritt vor Ende des betreffenden Heim- oder Spitalaufenthaltes, sonst nachher)
Wertebereich: -23'625 - 51'590

VONORG: kommt die Person aus einem Spital oder Heim

- 2 von Kurz-Spitalaufenthalt < 1 Tag
- 1 von Spitalaufenthalt > 0 Tage
- 0 weder von Spital noch Heim
- 1 von Heimaufenthalt > 0 Tage
- 2 von Kurz-Heimaufenthalt < 1 Tag

ZUORG: geht die Person in ein Spital oder Heim

- 2 zu Kurz-Spitalaufenthalt < 1 Tag
- 1 zu Spitalaufenthalt > 0 Tage
- 0 weder zu Spital noch Heim
- 1 zu Heimaufenthalt > 0 Tage
- 2 zu Kurz-Heimaufenthalt < 1 Tag

SEQ_U_all_neu: *Sequenznummer* (absolute Reihenfolge aller als valide eingeschätzten Spital- und Heimaufenthalte nach Eintrittsdatum)
alle Meldungen werden gezählt, ausser Fälle ohne Austritt, aber mit Folgeaufenthalt(en)
Wertebereich: 1 - 335

SPIOVER: *(Anzahl Heimaufenthaltstage jedes einzelnen Aufenthaltes, die durch Spitalaufenthalte überlappt werden)*
Wertebereich: -1146 - 20'305

SPIOVER2: *(Netto-Anzahl Heimaufenthaltstage, die durch einen Spitalaufenthalt überlappt werden, d.h. korrigiert für gleichzeitig überlappende Heimaufenthalte)*
Wertebereich: 0 - 20'305

7 Weitere Datenfehler

- (A) Mehrfache Todesmeldungen für denselben anonymen Verbindungscode (AVC)
- (B) Todesmeldungen für Personen mit weiteren Spital-/Heimaufenthalten nach dem Todesdatum (in der Regel Austritte, nicht selten aber auch Eintritte)
- (C) Spital-/Heimaufenthalte mit dem mutmasslichen Geburtsdatum als Eintrittsdatum.

7.1 Mehrfache Todesmeldungen für denselben AVC

Drei mögliche Ursachen für solche Meldungen sind denkbar:

1. Die Person ist bei der /den früheren Meldung(en) nicht gestorben
2. Der anonyme Verbindungscode ist fehlerhaft und die Todesmeldungen betreffen unterschiedliche Personen
3. Bei der späteren Meldung handelt es sich um eine verspätete Meldung (typisches Beispiel: Ein Heimklient muss wegen einer akuten Erkrankung ins Spital, tritt aber nicht aus dem Heim aus, weil er wieder zurückerwartet wird – er verstirbt aber im Spital und wird von diesem korrekt als gestorben gemeldet. Das Heim erfährt, dass der Patient verstorben ist und meldet als Austrittsgrund „gestorben“, mit dem Datum, an dem das Zimmer wieder für einen neuen Klienten zur Verfügung steht.

Insgesamt sind 23'374 AVCs von mehrfachen Todesmeldungen betroffen (Tab. 12).

Tab. 12 Anzahl AVCs mit mehrfachen Todesmeldungen im Zeitraum 2000-2016

Anzahl Todesmeldungen	Anzahl betroffene Verbindungscode
2	23'136
3	229
4	4
5	5
Total	23'374

Eine Durchsicht aller Fälle ist mit vernünftigem Zeitaufwand nicht möglich. Eine stichprobenartige Durchsicht zeigt aber häufige Muster auf. Zur Eruierung der jeweiligen Ursache wurde ein mehrstufiges Verfahren gewählt.

- Die neun AVCs mit **vier** respektive **fünf** Todesmeldungen wurden manuell bereinigt.
- Bei AVCs mit **drei** Todesmeldungen wurde
 - a) bei Personen mit mindestens einer Todesmeldung aus einem Spital das späteste Spital-Todesdatum akzeptiert, die anderen Todesmeldungen hingegen auf "nicht verstorben" umgesetzt. Wenn eine Meldung mit späterem Todesdatum aus einem Heim stammte, wurde zusätzlich das Austrittsdatum des Heims auf das späteste Todesdatum des Spitals zurückgesetzt
 - b) bei mehreren Todesmeldungen mit gleichem Austrittsdatum aus einem Spital wurde diejenige Meldung mit dem spätesten Eintrittsdatum als massgeblich akzeptiert
 - c) bei mehreren Todesmeldungen ausschliesslich aus Heimen wurde diejenige mit dem spätesten Eintrittsdatum als massgeblich akzeptiert.

- Bei **zwei** Todesmeldungen wurde wie folgt vorgegangen:
- d) Das mit Abstand häufigste Muster war: Heim und Spital melden einen Austritt mit Austrittsgrund „Tod“, das Heim jedoch zu einem späteren Zeitpunkt. Beträgt die Differenz maximal 31 Tage, wurde in allen Fällen die Heimmeldung als verspätete Todesmeldung interpretiert und die Heimaufenthaltsdauer auf das vom Spital gemeldete Todesdatum bezogen.
=> In 19'205 Fällen konnte die Todesmeldung auf den relevanten (überlappenden) Spitalaufenthalt beschränkt werden³. Es verbleiben 3'931 AVCs mit doppelten Todesmeldungen.
 - e) Der "umgekehrte" Fall: Das Heim meldet das korrekte Datum für den Übertritt ins Spital, aber nimmt den kurze Zeit später eintretenden Tod im Spital (spätestens 31 Tage später) in der Meldung (welche administrativ erst früh im Folgejahr erfolgt) als Austrittsgrund bereits vorweg:
=> 929 „verfrühte“ Todesmeldungen! Es verbleiben 3'002 AVCs mit doppelten Todesmeldungen
 - f) Wenn zwei Todesmeldungen aus Spitälern vorliegen, aber bei allen Spitalaufenthalten dieser Person (auch denjenigen ohne Todesmeldung) der Wohnkanton derselbe ist, wurde davon ausgegangen, dass die erste Todesmeldung falsch ist. Hospitalisierungen mit unterschiedlichem Wohnkanton traten nur bei 91 AVCs mit zwei Todesmeldungen aus Spitälern auf.
=> 1'036 erste Todesmeldungen verworfen. Es verbleiben 1'966 AVCs mit doppelten Todesmeldungen
 - g) Wenn zwei Todesmeldungen aus Heimen vorliegen und bei allen Aufenthalten dieser Person (also auch denjenigen ohne Todesmeldung) die ersten zwei Ziffern der vierstelligen Postleitzahl (Postleitzahlbereich) identisch sind, wurde davon ausgegangen, dass die erste Todesmeldung falsch ist. Von 1'261 AVCs mit zwei Todesmeldungen aus Heimen wiesen 1'089 immer dieselben Anfangsziffern bei der Postleitzahl auf.
=> 1'089 erste Todesmeldungen verworfen. Es verbleiben 877 AVCs mit doppelten Todesmeldungen
 - h) Die Bedingungen von f) und g) wurden insofern aufgeweicht, als Wohnkanton respektive zweistellige Postleitzahl nur noch bei den jeweiligen Todesmeldungen identisch sein mussten.
Bei jeweils zwei Spital- (respektive jeweils zwei Heim-) Todesmeldungen wurde diejenige mit dem früheren Austrittsdatum als "nicht gestorben" gewertet, bei gleichem Austrittsdatum die Meldung mit dem früheren Eintrittsdatum.
=> 139 falsche Todesmeldungen! Es verbleiben 738 AVCs mit doppelten Todesmeldungen
 - i) Die verbleibenden Meldungen, wo entweder der Wohnkanton oder der Postleitzahlbereich bei den Todesmeldungen nicht übereinstimmt, oder die Meldungen aus Spital und Heim stammen, wurden individuell durchgesehen und wie folgt gewertet:
=> 365 Fälle als verfrühte Todesmeldung
=> 222 Fälle als verspätete Meldung (die Aufenthaltsdauer wurde angepasst)
=> 151 Fälle als fehlerhafter/ungültiger Verbindungscode.

³ In 14'182 Fällen (73.8%) hatte die Todesmeldung des Heimes dasselbe Datum wie die des Spitals; die mittlere Verkürzung des Aufenthaltes bei den anderen 26.2% beträgt 9.50 Tage (bzw. 2.49 Tage bezogen auf alle SOMED-Aufenthalte); siehe auch Zellweger&Bopp 2017

7.2 Todesmeldung, aber trotzdem spätere Spital-/Heimaufenthalte

Bei insgesamt 80'272 Spital-/Heimaufenthalten von 9'356 Personen wurde 14'684 Mal ein späteres Austrittsdatum als das Todesdatum gemeldet:

späterer Austritt aus einem Heim	7'839
späterer Austritt aus einem Spital	6'845
Total	14'684

Ähnlich wie bei den mehrfachen Todesmeldungen wurde auch hier ein schrittweiser Lösungsansatz gewählt.

- a) Das mit Abstand häufigste Muster: Das Heim meldet einen Austritt 1 bis maximal 31 Tage nach dem Tod als Austritt; in der Regel als Übertritt in ein Spital: In all diesen Fällen wurde die Heimmeldung als verspätete Austrittsmeldung klassiert und die Heimaufenthaltsdauer auf den vermeldeten Todestag begrenzt.
=> In 6'060 Fällen konnte so die Austrittsmeldung und die damit verbundene Aufenthaltsdauer auf die relevante Todesmeldung reduziert werden. Es verbleiben 3'296 verschiedene AVCs mit Austrittsmeldungen nach dem Tod.
- b) Wurde ein Eintrittsdatum nach dem "Todesdatum" mit gleichem Wohn-Kanton bzw. gleichem Postleitzahlbereich gemeldet, wurde die Person als nicht gestorben klassifiziert.
=> In 1'729 Fällen wurde die Todesmeldung verworfen. Es verbleiben 1'567 verschiedene AVCs mit Austrittsmeldungen nach dem Tod.
- (c) Gab es spätere Eintritte mit anderem Wohnkanton oder anderem Postleitzahlbereich wurde primär ein falscher AVC vermutet. Eine stichprobenweise Überprüfung zeigte, dass dieses Muster nicht eindeutig ist: Hier wurden deshalb die Spital-/Heimaufenthalte für jeden der 275 betroffenen AVC einzeln beurteilt:
 - bei 708 Records von 264 verschiedenen AVCs wurde der AVC als falsch eingestuft (und dadurch 1621 Spital-/Heimaufenthalte von 265 Personen "gerettet")
 - 3 Fälle wurden als nicht verstorben klassiert (und dadurch 17 Spital-/Heimaufenthalte "gerettet")
 - 8 Fälle wurden als verspätete Meldung klassiert und das Austrittsdatum angepasst (und dadurch 101 Spital-/Heimaufenthalte von 8 Personen "gerettet")
=> Es verbleiben 1'292 verschiedene AVCs mit Austrittsmeldungen nach dem Tod.
- d) Bei Todes-Meldungen aus einem Spital, späteren Austrittsmeldungen aus einem Heim (oder umgekehrt) wurden die Fälle einzeln überprüft.
=> 69 Personen wurden als nicht gestorben klassiert;
=> In 628 Fällen wurde das Austrittsdatum korrigiert.
Es verbleiben 595 verschiedene AVCs mit Austrittsmeldungen nach dem Tod.
- e) Die verbliebenen 595 AVCs mit komplexen Meldungen wurden ebenfalls einzeln überprüft.
=> 71 AVCs wurde mindestens ein Record als falscher AVC klassifiziert
=> 232 Personen wurden als nicht gestorben interpretiert
=> 291 Fällen wurde das Heim-Austrittsdatum korrigiert.

Insgesamt (a bis e) wurden somit bei 344 verschiedenen AVCs insgesamt 914 Aufenthalte als ungültig deklariert, 2'033 Todesmeldungen als nicht verstorben klassiert (was insgesamt 6'783 Aufenthalte als korrekt deklariert) und 6'987 Aufenthalte als solche mit verspäteter Austrittsmeldung.

7.3 Verdächtig lange Aufenthaltsdauern (seit Geburt?)

Neben 58 Fällen ohne Altersangabe fallen sich über Jahrzehnte erstreckende Langzeitaufenthalte in Spitälern und Heimen auf. Erfahrungsgemäss werden Geburts- und Eintrittsdatum ab und zu verwechselt. Dazu kommen Tippfehler bei der Eingabe von Kalenderdaten. Da in der MedStat und in der SOMED das Geburtsdatum nur für Verstorbene verfügbar ist, lassen sich solche Verwechslungen nicht direkt sondern bloss über weitere Indizien (z.B. sehr lange Aufenthaltsdauern) aufdecken. Da häufig keine Gewissheit besteht, welches der richtige Wert sein müsste, beschränken wir uns auf die Definition einer Marker-Variablen GJ, haben aber die Originalangaben nicht verändert.

Beispiel 1: Verwechslung von Geburtsjahr und Eintrittsjahr

avc	invalid	jahr	betrieb	seq	sx	alter	eintritt	austritt	dauer	gj
xyz	späterer He	2007	A	.	Frau	87	26 Apr 20	.	32026	1
xyz	gueltig	2008	A	1	Frau	88	26 Apr 20	.	32392	1
xyz	gueltig	2009	A	2	Frau	89	19 Sep 03	08 Nov 09	2242	

2007 und 2008 wurde wohl das Geburtsdatum als Eintrittsdatum eingesetzt und dann im Jahr 2009 auf den tatsächlichen Wert korrigiert. Damit wird allerdings seq=1 zu Unrecht als gültig statt als redundant eingestuft.

Üblicherweise verbringen Menschen nicht ihr gesamtes Leben von Geburt an in ein- und derselben Institution. Deshalb sind Eintrittsdaten, die auf einen Aufenthalt seit Geburt hinweisen, primär etwas suspekt und lassen eine Verwechslung der Eingabefelder Geburtsdatum und Eintrittsdatum vermuten (ausgenommen davon sind natürlich Neugeborene im Spital). Trotzdem kann nicht ausgeschlossen werden, dass vereinzelte Person seit Geburt in einer Institution leben (müssen), z.B. wegen schweren Krankheiten oder Missbildungen. Eine individuelle Überprüfung ergab, dass in etwa jedem vierten Fall das Geburtsdatum tatsächlich das korrekte Eintrittsdatum sein könnte (z.B. Fälle mit Epilepsie, Zerebralparese oder andere Schwer(st)-behinderungen). Prinzipiell kommt jeder Spital- oder Heimaufenthalt mit Übereinstimmung von Eintrittsjahr und Geburtsjahr für eine Verwechslung in Frage und wurde daher in der Variable GJ mit dem Wert "1" markiert. Gerade in der MedStat ist dies bei Geburten allerdings der Normalfall.

Insgesamt wurden 226 erwachsene Personen mit mutmasslichem falschem Eintrittsdatum individuell überprüft. In 56 Fällen scheint das angegebene Datum wirklich zu stimmen; in 78 Fällen kann das korrekte Eintrittsdatum aufgrund anderer Spital- oder Heimaufenthalte eruiert werden; in 92 Fällen scheint schliesslich das Eintrittsdatum falsch zu sein, aber es gibt keine Hinweise auf ein vermutlich korrektes Eintrittsdatum.

Beispiel 2: Eintritts- oder Austrittsdatum fehlt und damit die Aufenthaltsdauer

avc	invalid	jahr	sx	alter	eintritt	austritt	dauer	gj
zzz	gueltig	2001	Frau	75-79
zzz	gueltig	2001	Frau	75-79	.	.	.	2
zzz	gueltig	2009	Frau	85-89	18 Jun 09	24 Jun 09	6	.

Die 218 Fälle in der Ursprungsversion bzw. 173 gültigen Hospitalisierungen ohne genügende Angaben für die Berechnung der Aufenthaltsdauer (fast alle davon vor 2009) können nicht befriedigend gelöst werden, wurden aber alle in der Variable GJ mit dem Wert "2" markiert.

Beispiel 3: Falsches Jahrhundert beim Eintrittsjahr (z.B. 1904 statt 2004)

avc	invalid	jahr	sx	eintritt	austritt	alter	dauer	gj
yxa	spaeterer He	2007	Frau	04 Jun 1904	.	96	37831	3
yxa	queltig	2008	Frau	04 Jun 1904	24 Mar 08	97	37914	3

Bei 25 SOMED-Fällen (davon 16 "gültigen") scheint dieser Fehler eindeutig. Diese Fälle wurden daher in der Variable GJ mit dem Wert "3" markiert.

Beispiel 4: Schreibfehler

avc	invalid	jahr	sx	eintritt	austritt	alter	dauer	gj
cba	spaeterer He	2007	Mann	29 Oct 70	.	27	13578	4
cba	gueltig	2008	Mann	29 Oct 70	17 Mar 08	28	13654	4

Der Schreibfehler scheint in diesem Fall offensichtlich, weil der Patient 1970 noch gar nicht geboren war – richtig ist wohl 2007 statt 1970. Insgesamt fallen 51 SOMED-Fälle (davon 32 "gültige") in diese Kategorie; alle wurden in der Variable GJ mit dem Wert "4" markiert.

GJ potentieller Fehler im Eintritts- oder Geburtsdatum

- 1 Geburtsjahr = Eintrittsjahr (ohne besondere Argumente für eine Verwechslung)
- 2 Angabe der Aufenthaltsdauer fehlt
- 3 Eintrittsjahr: Jahrhundert falsch? (nur SOMED)
- 4 Tippfehler bei Datum? (nur SOMED)

8 Herkunft und Destination bei Spital- und Heimaufenthalten

Stichproben haben gezeigt, dass die Angaben, von woher die Patienten ins Spital oder Heim kommen bzw. wohin sie nach dem Spital-/Heimaufenthalt entlassen werden, nicht immer zuverlässig sind.

Es wurden deshalb zusätzliche Variablen definiert, die zumindest direkt vorgängige und anschliessende Spital- und Heimaufenthalte aufzeigen (anschliessende Heimaufenthalten wurden auch dann als solche gekennzeichnet wenn sie innerhalb von 30 Tagen nach Spitalaustritt erfolgten). In den Tabellen 13 und 14 zeigen grau markierte Flächen die Übereinstimmung der Ursprungsangaben in MedStat bzw. SOMED mit den neu gebildeten Variablen.

8.1 Bewegung zu/von Spitälern

Im Folgenden wird die Übereinstimmung der ursprünglichen MedStat-Angaben zu Herkunft und Destination mit den aus den kombinierten Daten selbst herstellbaren Wegen verglichen. Mit 93.6%⁴ Übereinstimmung bezogen auf die drei Kategorien Spital, Heim und "Anderes" (Tab. 13a) erscheint sie auf den ersten Blick relativ hoch, was jedoch darauf zurückzuführen ist, dass die allermeisten Spitaleintritte von Zuhause aus erfolgen.

Betrachtet man die Übereinstimmung bei der Herkunft aus einem anderen Spital oder aus einem Heim, liegt sie wesentlich tiefer. Nur bei 161'000 Aufenthalten stimmt die Herkunft Heim überein, bei knapp 200'000 Aufenthalten finden wir dagegen keinen vorgängigen Heimaufenthalt in der SOMED. Andererseits zeigt die Datenkombination in knapp 280'000 Fällen, dass die Patienten direkt aus einem Heim ins Spital eintraten, obwohl in der MedStat ein anderer Herkunftsort ausgewiesen wird (250'000 Mal von zu Hause).

Ein ähnliches Bild zeigt sich bei den Spitalaustritten, mit einer globalen Übereinstimmung von 93.2%⁵ Übereinstimmung (Tab. 13b), grossmehrheitlich zurückzuführen auf die Entlassung nach Hause. Dagegen stimmen nur etwa ein Drittel der Fälle mit Entlassungsdestinationen "anderes Spital" und bloss rund die Hälfte mit der Entlassungsdestination "Heim" überein.

Tab. 13a Herkunft bei Spitalaufenthalten

(Originalvariable in Spalte A, rekonstruierte Variable in Zeile A)

A	Spital	keine Org	Heim	Total
Zu Hause	259'753	17'577'499	251'672	18'088'924
SOMED *	27'919	169'848	161'028	358'795
Spital**	1'237'856	541'929	11'409	1'791'194
Anderere	64'519	503'207	10'000	577'726
Unbekannt	21'920	333'941	5'164	361'025
Total	1'611'967	19'126'424	439'273	21'177'664
Übereinstimmung		19'813'531	93.6%	21'177'664

* Altersheime, Pflegeheime, Institutionen für Behinderte, Institutionen für Suchtkranke, Institutionen für Personen mit psychosozialen Problemen.

** Zentrumsversorgung, Grundversorgung, Psychiatrische Kliniken, Rehabilitationskliniken und andere Spezialkliniken

⁴ 19'813'531 / 21'177'664 Eintritte

⁵ 19'349'719 / 20'759'766 Austritte

Tab. 13b Destination nach Spitalaufenthalten

(Originalvariable in Spalte A, konstruierte Variable in Zeile A)

A	Spital	keine Org	Heim	Total
Verstorben	1'211	1'183	62	2'456
Nach Hause	172'746	17'015'439	198'376	17'386'561
SOMED *	44'119	356'969	392'446	793'534
Spital**	1'349'387	511'281	25'715	1'886'383
Andere	40'862	146'888	18'172	205'922
Unbekannt	29'210	445'559	10'141	484'910
Total	1'637'535	18'477'319	644'912	20'759'766
Übereinstimmung		19'349'719	93.2%	20'759'766

* Altersheime, Pflegeheime, Institutionen für Behinderte, Institutionen für Suchtkranke, Institutionen für Personen mit psychosozialen Problemen.

** Zentrumsversorgung, Grundversorgung, Psychiatrische Kliniken, Rehabilitationskliniken und andere Spezialkliniken

8.2 Bewegung zu/von Heimen

Bei den Heimklienten ist die Übereinstimmung bezüglich der Herkunft mit nur 61%⁶ deutlich tiefer als in der MedStat (Tabelle 14a). Dies ist darauf zurückzuführen, dass der Eintritt von "zu Hause" deutlich seltener ist als bei einem Spitalaufenthalt. Die knapp 300'000 Eintritte, die übereinstimmend als nicht aus einem Spital oder Heim stammend ausgewiesen werden, entsprechen knapp 80% der Eintritte, die die SOMED als solche aufführt bzw. knapp 60% der Eintritte, wie wir sie berechnet haben. 21'000 Übertritte von einem Heim in ein nächstes stimmen überein, aus Sicht der SOMED sind das knapp 30%, aus unserer Sicht knapp ein Drittel. Bei einem Heim-Eintritt aus einem Spital ist die Übereinstimmung wesentlich besser: Bei rund zwei Dritteln der Aufenthalte stimmen diese Angaben überein.

Bei den Heimaustritten ist die globale Übereinstimmung mit 70%⁷ besser (Tabelle 14b). Aber auch in den Details ist die Übereinstimmung besser: Heimaustritte "nach Hause" stimmen bei Klienten, die nicht verstorben sind, zu rund 85% überein. Der Übertritt in ein weiteres Heim stimmt aus Sicht der SOMED zu zwei Dritteln, aus Sicht unserer rekonstruierten Daten zu knapp der Hälfte aller Fälle überein; der Übertritt von einem Heim in ein Spital zu 57% respektive 46%.

Tab. 14a Herkunft bei Heimaufenthalten

(Originalvariable in Spalte A, konstruierte Variable in Zeile A)

A	Spital	keine Org	Heim	Total
Zu Hause	52'356	259'420	20'536	332'312
SOMED *	19'440	33'303	20'924	73'667
Spital**	173'292	76'172	7'424	256'888
Andere	9'771	23'054	8'470	41'295
Unbekannt	4'152	12'948	1'031	18'131
Total	263'852	476'570	64'209	804'631
Übereinstimmung		489'638	60.9%	804'631

* Altersheime, Pflegeheime, Institutionen für Behinderte, Institutionen für Suchtkranke, Institutionen für Personen mit psychosozialen Problemen.

** Zentrumsversorgung, Grundversorgung, Psychiatrische Kliniken, Rehabilitationskliniken und andere Spezialkliniken

⁶ 489'638 / 804'631 Heimeintritte

⁷ 240'526 / 333'319 Heimaustritte

Tab. 14b Destination nach Heimaufenthalten

(Originalvariable in Spalte A, konstruierte Variable in Zeile A)

A	Spital	keine Org	Heim	Total
Verstorben	3'091	18'880	1'052	23'023
Nach Hause	4'245	163'425	18'426	186'096
SOMED *	4'077	12'830	34'808	51'715
Spital**	12'723	7'131	2'328	22'182
Andere	2'546	18'446	15'502	36'494
Unbekannt	702	11'124	1'983	13'809
Total	27'384	231'836	74'099	333'319
Übereinstimmung		240'526	69.6%	333'319

* Altersheime, Pflegeheime, Institutionen für Behinderte, Institutionen für Suchtkranke, Institutionen für Personen mit psychosozialen Problemen.

** Zentrumsversorgung, Grundversorgung, Psychiatrische Kliniken, Rehabilitationskliniken und andere Spezialkliniken

9 Fazit

Ungültige Aufenthaltsmeldungen und falsche Todesmeldungen lassen sich auf Datenfehler in drei Kategorien zurückführen:

- falsche anonyme VerbindungsCodes und Mehrfachmeldungen eines Aufenthaltes führen zu ungültigen Aufenthaltsmeldungen
- fehlerhafte Angaben zum Tod (mehrere Institutionen oder Abteilungen derselben Institution melden den Tod, falsch datierte oder unzutreffende Todesmeldungen, wenn die Person überlebt hat).

Zum Teil als Folge dieser Datenfehler enthalten die Daten weitere Fehler, die soweit möglich behoben werden sollten, insbesondere bei den Sequenznummern und den Angaben zu Herkunft vor und Destination nach dem Spital-/Heim-Aufenthalt.

9.1 Ungültige Fälle

Über alles betrachtet (Tabelle 15), müssen 14.6% der Aufenthaltsmeldungen ("Fälle") (MedStat 2000-2016: 11%, SOMED-Daten 2007-2016: 59%) für die Analyse der Spital-/Heimaufenthalte auf Personenebene ausgeschlossen (bzw. als ungültig markiert) werden, da sie

- a) nur lückenhaft erhoben worden sind: Ambulante und teilstationäre Hospitalisierungen 2000-2008
- b) redundant sind (B-/C-Fälle der MedStat und SOMED-Aufenthalte mit Fortsetzungsrecord im Folgejahr)
- c) fehlerhaft sind (Mehrfachmeldungen, falscher/ungültiger AVC).

Lückenhaft erhobene und redundante Meldungen machen die überwiegende Mehrzahl der ausgeschlossenen Fälle aus. Sie sind in der Regel einfach zu erkennen und herauszufiltern, bei den redundanten Fällen stimmen dann aber manchmal nicht alle Angaben mit denen des Folgejahres überein.

Tab. 15 Als ungültig erachtete Fälle

Status	Total	Spital	Heim
Total	25'744'296	23'788'383	1'955'913
gültig	21'982'295	21'177'664	804'631
ungültig:	3'762'001	2'610'719	1'151'282
– mit Sicherheit ungültig bzw. redundant:	3'638'629	2'487'699	1'150'930
• AVC fehlt	118'737	118'347	390
• B-/C-Fall mit späterem A-Fall	315'161	315'161	
• ambulante oder teilstationäre Hospitalisierung	2'054'191	2'054'191	
• Heimrecord in einem Folgejahr mit demselben Eintrittsdatum	1'150'540		1'150'540
– Datenfehler:	123'372	123'020	352
• sichere/wahrscheinliche Mehrfachmeldung	13'966	13'966	
• Ungültiger AVC	109'276	109'054	222
• andere Gründe*	130		130

* für die meisten Fälle wurde im Folgejahr ein Record gefunden, der zwar nicht das identische - aber ein plausibles - Eintrittsdatum aufweist: dieser Record wurde deshalb als Fortsetzungsrecord interpretiert und der Vorjahresrecord als redundant.

Tab. 16a Als ungültig erachtete Spitalaufenthalte nach Kalenderjahr

	Total	mit Sicherheit ungültig		potenziell gültig	"wahrscheinlicher Datenfehler"				gültig
		AVC fehlt	B/C-Fall mit folgendem A-Fall		AVC invalid	Subtotal	N	% aller potenziell gültigen Fälle	
2000	1'184'229	65'329	15'188	165'511	1'384	2'817	4'201	0.45%	934'000
2001	1'302'964	33'536	17'627	220'680	1'164	4'505	5'669	0.55%	1'025'452
2002	1'346'129	17'208	19'198	199'153	2'499	1'342	3'841	0.35%	1'106'729
2003	1'390'903	98	20'990	220'780	5'199	913	6'112	0.53%	1'142'923
2004	1'409'743	106	19'957	213'039	539	1'183	1'722	0.15%	1'174'919
2005	1'466'467	158	20'648	248'127	261	38'971	39'232	3.28%	1'158'302
2006	1'500'664	127	20'663	261'300	1'027	2'649	3'676	0.30%	1'214'898
2007	1'538'323	114	19'596	258'379	187	3'635	3'822	0.30%	1'256'412
2008	1'584'754	1'237	19'390	267'222	483	30'058	30'541	2.35%	1'266'364
2009	1'328'127	72	19'861	0	339	2'476	2'815	0.22%	1'305'379
2010	1'345'245	61	20'342	0	200	83	283	0.02%	1'324'559
2011	1'363'698	41	17'680	0	201	20'086	20'287	1.51%	1'325'690
2012	1'353'521	182	19'969	0	84	59	143	0.01%	1'333'227
2013	1'374'454	25	21'634	0	218	57	275	0.02%	1'352'520
2014	1'401'014	13	20'301	0	70	58	128	0.01%	1'380'572
2015	1'430'201	17	22'117	0	34	99	133	0.01%	1'407'934
2016	1'467'947	23	0	0	77	63	140	0.01%	1'467'784
Total	23'788'383	118'347	315'161	2'054'191	13'966	109'054	123'020	0.58%	21'177'664

Tab 16b Als ungültig erachtete Heimaufenthalte nach Kalenderjahr

	Total	AVC fehlt	Späterer Heimrecord mit selbem Eintritt	potenziell gültig	"wahrscheinlicher Datenfehler"				gültig
					AVC invalid	anderes gültiges Eintrittsdatum	falscher AVC aufgrund Hospitalisierung nach Tod erkannt	Invalid aufgrund negativer, korrigierter Dauer	
2007	183'631	309	121'492	61'830	87	0	3	7	61'733
2008	186'328	81	123'495	62'752	9	37	6	3	62'697
2009	190'352	0	126'100	64'252	11	18	2	3	64'218
2010	189'117	0	125'964	63'153	9	17	3	6	63'118
2011	190'692	0	127'519	63'173	13	1	5	5	63'149
2012	195'931	0	130'538	65'393	4	0	2	5	65'382
2013	200'875	0	132'730	68'145	6	1	4	6	68'128
2014	204'445	0	133'457	70'988	13	0	6	5	70'964
2015	211'646	0	129'245	82'401	6	0	7	7	82'381
2016	202'896	0	0	202'896	18	1	8	8	202'861
Total	1'955'913	390	1'150'540	804'983	176	75	46	55	804'631

Nur ein kleiner Teil der von uns als ungültig markierten Fälle geht auf eigentliche Datenfehler zurück (praktisch Null in der SOMED; 0.6% aller potenziell gültigen Fälle der MedStat) (Tab. 16a), wobei die Datenqualität über die Jahre zunimmt. Ab 2010 (mit Ausnahme des Jahres 2011) beträgt der Anteil der ausgeschlossenen an allen potenziell gültigen Fällen maximal 0.2 Promille.

9.2 Falsche Todesmeldung

Falsche Todesmeldungen sind in der MedStat (Tabellen 17a/18a) relativ selten (0.57% aller Todesmeldungen) und sie nehmen über die Zeit markant ab (von 2% im Jahre 2000 bis 0.1% 2016).

In der SOMED (Tab. 17b/18b) hingegen ist dieser Anteil markant höher: den 241'000 akzeptierten Todesmeldungen stehen 23'000 nicht akzeptierte gegenüber (9.6%). Dieser Anteil ist über die Jahre sehr konstant (9.0 bis 10.1%) und in der Tendenz sogar leicht ansteigend.

Der Grossteil der falschen Todesmeldungen ist einfach zu erklären: In 19'580 Fällen ist die Person maximal 31 Tage vorher in einem Spital verstorben. Man kann hier den Todesort "Spital" als den richtigen akzeptieren, da die SOMED häufig nicht das eigentliche Todesdatum meldet, sondern das Datum, ab dem das Zimmer wieder für neue Klienten zur Verfügung steht. Da SOMED-Daten immer anfangs des Folgejahres erfasst werden, kann man sich leicht vorstellen, dass bei Klienten, die den Platz nicht gekündigt hatten, sondern wegen einer akuten Erkrankung hospitalisiert wurden und dabei gestorben sind, als Austrittsgrund (der bekannte) Tod und als Austrittsdatum dasjenige des freigewordenen Zimmers gemeldet wurden.

Tab. 17a Entscheid für Austritt aus Spital ("Austrittsgrund")

	Neu	Ursprünglich
nicht gestorben, da ein Aufenthalt mit späterem Austrittsdatum vorhanden ist*	12	
späterer Eintritt mit gleichem Wohnkanton	1'130	
später gestorben	141	
es gibt einen Eintritt nach der Todesmeldung	34	
bei zwei Todesmeldungen mit identischem Wohnkanton die frühere	1'036	
bei mehr als zwei (?) Todesmeldungen diejenige mit dem frühestem Eintrittsdatum*	90	
Total nicht bei diesem gemeldeten Aufenthalt gestorben:	2'443	
1 Initiative des Behandelnden	21'390'976	21'390'976
2 Initiative des Patienten	466'255	466'255
3 Initiative Drittperson	51'762	51'762
4 interner Übertritt/Wechsel der Hauptdiagnose	111'847	111'847
5 Gestorben	425'251	427'694
8 Anderes	244'653	244'653
9 Unbekannt	1'095'196	1'095'196
Total	23'788'383	23'788'383

* aus manuellem Vergleich aller Records dieses Patienten

Tab. 17b Falsche Todesmeldung/Datenfehler in SOMED

Austritt aus Heim wohin:	Neu	Ursprünglich
-11 nicht gestorben wegen späteren Daten	217	
-6 späterer Eintritt mit identischer Postleitzahlregion	602	
-5 später gestorben	512	
-4 Eintritt nach Tod => nicht gestorben	108	
-3 zwei Mal im Heim verstorben, immer Postleitzahlregion identisch	1'089	
-2 von Hand: hier nicht gestorben, bei mehreren Toden	947	
-1 bei mehreren Todesmeldungen diejenige des Spitaltodesdatum, bei gleichem Datum, dasjenige mit niedrigsten Eintrittsdatum	19'580	
Total nicht bei diesem gemeldeten Aufenthalt gestorben:	23'055	
1 nach Hause	207'694	207'694
2 Somed*	55'852	55'852
3 Spital**	24'199	24'199
4 Andere Institution	11'837	11'837
5 Anderes	29'497	29'497
6 Verstorben	262'670	285'725
99 Unbekannt	17'548	17'548
Total	632'352	632'352
Meldungen ohne Austritt (Ende Meldejahr noch im Heim)	1'500'628	1'500'628
	2'132'980	2'132'980

* Altersheime, Pflegeheime, Institutionen für Behinderte, Institutionen für Suchtkranke, Institutionen für Personen mit psychosozialen Problemen.

** Zentrumsversorgung, Grundversorgung, Psychiatrische Kliniken, Rehabilitationskliniken, Andere Spezialkliniken

Tab. 18 "Todesfälle" im Spital (oben) und Heim (unten), nach Kalenderjahr

	Vom Spital als verstorben gemeldet	als nicht (bei diesem Aufenthalt) verstorben beurteilt:						Als ver- storben akzeptiert
		A	B	C	D	E	F	
		Total						%
		N						
2000	21'915	0	314	13	1	109	2	2.00%
2001	23'525	1	208	9	1	89	8	1.34%
2002	24'521	2	116	10	1	92	11	0.95%
2003	24'852	1	76	19	0	86	5	0.75%
2004	24'212	0	63	7	0	150	3	0.92%
2005	24'923	0	79	16	2	111	1	0.84%
2006	24'699	0	69	15	4	69	4	0.65%
2007	25'332	1	30	6	2	63	8	0.43%
2008	25'466	2	17	8	3	41	12	0.33%
2009	25'778	1	28	13	2	37	4	0.33%
2010	25'617	1	23	5	0	31	6	0.26%
2011	25'639	2	36	4	3	40	4	0.35%
2012	25'897	0	27	5	5	36	9	0.32%
2013	26'558	1	19	7	2	25	6	0.23%
2014	26'024	0	10	2	1	20	6	0.15%
2015	26'584	0	8	1	4	20	1	0.13%
2016	26'152	0	7	1	3	17	0	0.11%
	427'694	12	1'130	141	34	1'036	90	0.57%
								425'251

Spaltenbedeutungen

- A** Manuell: nicht gestorben da späterer Austritt vorhanden
- B** es gibt späteren Eintritt mit gleichem Wohnkanton
- C** später gestorben
- D** es gibt Eintritt nach Todesmeldung
- E** bei mehreren Todesmeldungen mit identischem Wohnkanton immer die frühere
- F** Manuell: bei mehreren Todesmeldungen diejenige mit frühestem Eintrittsdatum

	Vom Heim als verstorben gemeldet	als nicht (bei diesem Aufenthalt) verstorben beurteilt											Als verstorben akzeptiert
		G	H	I	J	K	L	M	Total				
									N	%			
2007	23'337	18	40	53	13	131	96	1'676	2'027	9.51%	21'310		
2008	24'316	31	48	69	15	145	82	1'611	2'001	8.97%	22'315		
2009	25'088	23	41	57	12	135	83	1'957	2'308	10.13%	22'780		
2010	25'470	22	68	55	17	136	87	1'828	2'213	9.52%	23'257		
2011	25'518	15	37	62	12	104	97	1'789	2'116	9.04%	23'402		
2012	27'083	23	57	46	11	106	90	1'897	2'230	8.97%	24'853		
2013	27'624	31	73	42	11	106	111	2'065	2'439	9.68%	25'185		
2014	27'214	22	75	41	8	106	107	2'062	2'421	9.76%	24'793		
2015	30'039	25	112	49	5	74	98	2'371	2'734	10.01%	27'305		
2016	28'040	7	51	38	4	46	96	2'324	2'566	10.07%	25'474		
Total	263'729	217	602	512	108	1'089	947	19'580	23'055	9.58%	240'674		

- G** Manuell: nicht gestorben wegen späteren Daten
- H** bei späteren Eintritt mit identischer Postleitzahl-region
- I** später gestorben, Eintritt nach Tod
- J** nicht gestorben
- K** zwei Mal im Heim verstorben, immer Postleitzahlregion identisch
- L** manuell: hier nicht gestorben, bei mehreren Tode
- M** bei mehreren Todesmeldungen ist diejenige des Spitals gültig, bei gleichem Datum, diejenige mit höchstem Eintrittsdatum

9.3 Fehlerhafte Daten insgesamt

Insgesamt sind 108'500 (1.255% aller vorkommenden AVCs, die mindestens einen stationären Spitalaufenthalt oder einen Heimaufenthalt haben; vgl. Tab. 19) von erkennbaren Datenfehlern betroffen, sind also AVCs, die entweder offensichtlich nicht korrekt sind, oder AVCs, die mehrfache, verfrühte, verspätete oder falsche Todesmeldungen enthalten, oder mehrfache Aufenthaltsmeldungen des gleichen Aufenthaltes.

Die Probleme einzelner Spitäler in einzelnen Jahren bei der Zuweisung eines AVC hat beträchtliche jährliche Unterschiede beim Anteil der fehlerhaften AVCs zur Folge, nämlich zwischen einem Minimum von 0.187% im Jahr 2004 und einem Maximum von 3.116% im Jahr 2005. Seit 2010 ist die Fehlerquote mit Ausnahme des Jahres 2011 (1.532%) relativ niedrig, doch lässt sich seither kein klarer Trend zu einer weiteren Verbesserung feststellen.

Tab. 19 Fehlerhafte Verbindungscode nach Jahr

Jahr	AVCs, die in diesem Jahr vorkommen	davon AVCs, die in diesem Jahr mindestens einen fehlerhaften Record haben	
		N	%
2000	729'923	4'000	0.548%
2001	790'072	4'047	0.512%
2002	844'039	3'674	0.435%
2003	866'191	5'373	0.620%
2004	888'037	1'665	0.187%
2005	894'166	27'866	3.116%
2006	899'411	3'392	0.377%
2007	1'047'479	5'016	0.479%
2008	1'077'692	26'447	2.454%
2009	1'076'254	4'976	0.462%
2010	1'096'680	2'935	0.268%
2011	1'104'459	16'921	1.532%
2012	1'134'399	2'814	0.248%
2013	1'140'534	3'125	0.274%
2014	1'160'660	2'999	0.258%
2015	1'179'550	3'381	0.287%
2016	1'175'899	3'644	0.310%
Insgesamt	8'646'972	108'500	1.255%

10 Liste der neu gebildeten Variablen und ihre Kennzahlen

Nachstehend ein Überblick über die neu gebildeten Variablen (in Klammer Angabe des Kapitels in diesem Bericht, in dem die Variable eingeführt und beschrieben wird).

Variablenname	Bedeutung	Vorkommen	
		MedStat	SOMED
a) Bedeutung für Spital und Heimaufenthalte			
dauer_u	Aufenthaltsdauer in Tagen	(3)	(5.3)
seq_u	Sequenznummer MedStat/SOMED separat	(3)	(5.3)
seq_u_all	Sequenznummer MedStat/SOMED zusammen	(6.2)	(6.2)
overlap	Hospitalisierung ist überlappend	(3)	(5.3)
overlaprest	Anzahl Tage, die nicht von der vorhergehenden Hospitalisierung abgedeckt sind	(3)	(5.3)
tage_bis_last_hosp	Anzahl Tage bis zum letzten dokumentierten Spitalaufenthalt	(6.2)	(6.2)
invalid	Zusammenfassung der Gründe für Ungültigkeitserklärung	(6.2)	(6.2)
vonorg	kommt die Person aus einem Spital oder Heim	(6.2)	(6.2)
zuorg	geht die Person in ein Spital oder Heim	(6.2)	(6.2)
gj	potentielle Fehler im Eintrittsdatum	(7.3)	(7.3)
b) Bedeutung nur bei Spitalaufenthalten			
inv_vc	AVC ist ungültig	(2)	
inv_vc_perc	Anteil systematisch falscher AVCs, nach Spital und Statistikjahr	(2.1)	
folgehosp	Fortsetzungshospitalisierung prospektiv (Bedeutung nur für B-/C-Records)	(2.2)	
doppel	Potentielle Mehrfachmeldung	(2.3)	
seitlasthosp	Tage seit letztem Spitalaustritt	(3)	
samespit	gleiches oder anderes Spital	(3)	
lasthosp	Hospitalisierung mit spätestem Austrittsdatum	(3)	
c) Bedeutung nur bei Heimaufenthalten			
lastheim	Letzter Heimaufenthalt		(5.3)
seitlastheim	Tage seit letztem Heimaufenthalt		(5.3)
folgeheim	Fortsetzungsrecord im Folgejahr		(5.3)
spiover	Anzahl Heimaufenthaltstage, die mit Spitalaufenthalt(en) überlappen (brutto)		(6.2)
spiover2	Anzahl Heimaufenthaltstage, die mit Spitalaufenthalt(en) überlappen (netto, d.h. korrigiert für überlappende Heimaufenthalte)		(6.2)

Kennzahlen (MedStat 2000-16; SOMED 2007-16)

	alle Aufenthalte				alle gültigen Aufenthalte			
	Valid	Missing	Minimum	Maximum	Valid	Missing	Minimum	Maximum
invalid	25'744'296	0	0	12	21'982'295	0	0	0
inv_vc invalid AVC	23'823'175	1'921'121	0	6	21'212'104	770'191	0	2
inv_vc_perc % syst. inv. AVCs in Betrieb & Sta_jahr	23'788'383	1'955'913	0	1	21'177'664	804'631	0	1
folgehosp hat B/C-Fall eine Folge-Hospitalisierung?	23'788'383	1'955'913	-1	3	21'177'664	804'631	-1	0
doppel potentielle Doppelmeldung	23'788'383	1'955'913	0	11	21'177'664	804'631	0	11
folgeheim	1'955'913	23'788'383	-1	9	804'631	21'177'664	-1	2
seq_u (Reihenfolge Eintritt: Spital/Heim je separat)	21'983'857	3'760'439	1	334	21'982'295	0	1	334
seq_u_all_neu_ (Reihenfolge Eintritt Spital Heim zusammen)	25'744'296	0	1	118'737	21'982'295	0	1	335
dauer_u Austritt-Eintritt in Tagen	25'743'997	299	-30	154'113	21'982'050	245	0	154'113
vonorg	21'983'852	3'760'444	-2	2	21'982'295	0	-2	2
zuorg	21'983'852	3'760'444	-2	2	21'982'295	0	-2	2
tage_bis_last_hosp Tage bis Eintritt last Hosp. 2000-2016	21'835'118	3'909'178	-23'625	51'590	21'833'663	148'632	-23'625	51'590
samespit gleiches oder anderes Spital wie letzte H	23'769'684	1'974'612	0	1	21'169'904	812'391	0	1
seitlasthosp tage Eintritt minus Austritt last hosp	14'349'285	11'395'011	-19'082	6'197	13'098'022	8'884'273	-15'020	6'197
seitlastheim Eintritt minus Austritt last Heim	225'940	25'518'356	-27'981	3'628	225'860	21'756'435	-27'981	3'628
lasthosp Hospitalisierung mit latest discharge	9'337'070	16'407'226	0	2	8'561'786	13'420'509	0	2
lastheim Heimaufenthalt mit latest discharge	1'955'913	23'788'383	0	1	804'631	21'177'664	0	1
overlap überlappende hosp	25'744'296	0	0	2	21'982'295	0	0	2
overlaprest Dauer ohne overlaptage vorangehender hosp	25'605'419	138'877	-18	154'113	21'982'050	245	0	154'113
spiover	804'983	24'939'313	-1'146	20'305	804'631	21'177'664	-1'146	20'305
spiover2	21'985'075	3'759'221	0	20'305	21'982'295	0	0	20'305
gj	1'740'360	24'003'936	1	4	1'672'216	20'310'079	1	4
				gültige Spitalaufenthalte	gültige Heimaufenthalte			
invalid	21'177'664	0	0	0	804'631	0	0	0
inv_vc invalid AVC	21'177'664	0	0	2	34'440	770'191	0	0
inv_vc_perc % syst. inv. AVCs in Betrieb & Sta_jahr	21'177'664	0	0	1	0	804'631		
folgehosp hat B/C-Fall eine Folge-Hospitalisierung?	21'177'664	0	-1	0	0	804'631		
doppel potentielle Doppelmeldung	21'177'664	0	0	11	0	804'631		
folgeheim	0	21'177'664			804'631	0	-1	2
seq_u (Reihenfolge Eintritt: Spital/Heim je separat)	21'177'664	0	1	334	804'631	0	1	127
seq_u_all (Reihenfolge Eintritt Spital Heim zusammen)	21'177'664	0	1	335	804'631	0	1	281
dauer_u Austritt-Eintritt in Tagen	21'177'424	240	0	154'113	804'626	5	0	54'292
vonorg	21'177'664	0	-2	2	804'631	0	-2	2
zuorg	21'177'664	0	-2	2	804'631	0	-2	2
tage_bis_last_hosp Tage bis Eintritt last Hosp. 2000-2016	21'177'491	173	0	32'343	656'172	148'459	-23'625	51'590
samespit gleiches oder anderes Spital wie letzte H	21'169'904	7'760	0	1	0	804'631		
seitlasthosp tage Eintritt minus Austritt last hosp	13'098'022	8'079'642	-15'020	6'197	0	804'631		
seitlastheim Eintritt minus Austritt last Heim	0	21'177'664			225'860	578'771	-27'981	3'628
lasthosp Hospitalisierung mit latest discharge	8'561'786	12'615'878	0	2	0	804'631		
lastheim Heimaufenthalt mit latest discharge	0	21'177'664	0		804'631	0	0	1
overlap überlappende hosp	21'177'664	0	0	2	804'631	0	0	2
overlaprest Dauer ohne overlaptage von vorangehender hosp	21'177'424	240	0	154'113	804'626	5	0	54'292
spiover	0	21'177'664	0		804'631	0	-1'146	20'305
spiover2	21'177'664	0	0	0	804'631	0	0	20'305
gj	1'670'717	19'506'947	1	2	1'499	803'132	1	4

Literatur

- Bähler C, Signorell A, Reich O. Health care utilisation and transitions between health care settings in the last 6 months of life in Switzerland. *PLoS ONE* 2016;11(9):e0160932
- Berlin C, Busato A, Rosemann T, Djalali S, Maessen M. Avoidable hospitalizations in Switzerland: a small area analysis on regional variation, density of physicians, hospital supply and rurality. *BMC Health Services Res* 2014;14:289
- BFS 2009. 80 Tage im Spital vor dem Heimeintritt. Statistiken der stationären Gesundheitsversorgung – Studie zu den letzten Lebensjahren in Heim und Spital. Medienmitteilung 27.9.2009. <https://www.bfs.admin.ch/bfs/de/home/aktuell/neue-veroeffentlichungen.assetdetail.30501.html> (letzter Aufruf am 31. Mai 2019)
- BFS 2018. Wichtige Informationen zum Bezug von anonymisierten Einzeldaten aus der Medizinischen Statistik der Krankenhäuser der Sektion Gesundheit GES. <https://www.bfs.admin.ch/bfsstatic/dam/assets/319559/master> (letzter Aufruf am 31. Mai 2019)
- Creiser C. Erste Erhebung "Strukturdaten der Arztpraxen und ambulanten Zentren" (MAS 2015): Analyse von Teilnahme und Grundgesamtheit. Bundesamt für Statistik, Abteilung Gesundheit und Soziales, 5. Februar 2018. Neuchâtel: Bundesamt für Statistik
- Frick U, Frick H. «Heavy Use» in der stationären Psychiatrie der Schweiz? Ergebnisse aus der Medizinischen Statistik der Krankenhäuser. Obsan Dossier 11. Neuchâtel: Schweizerisches Gesundheitsobservatorium, 2010
- Frick U, Frick H. «Drehtüre» in der stationären Psychiatrie der Schweiz? Mythos oder empirische Realität? Obsan Dossier 12. Neuchâtel: Schweizerisches Gesundheitsobservatorium, 2010
- Füglister-Dousse S, Pellegrini S. Aufenthaltsorte von älteren Menschen im letzten Lebensabschnitt – Analyse der Abfolge von Aufhalten in Spital und Pflegeheim. Obsan Bulletin 2/2019. Neuchâtel: Schweizerisches Gesundheitsobservatorium
- Hedinger D, Braun J, Kaplan V, Bopp M. Determinants of aggregate length of hospital stay in the last year of life in Switzerland. *BMC Health Services Research* 2016;16:463
- Hedinger D, Haemmig O, Braun J, Bopp M. Social determinants of duration of last nursing home stay at the end of life in Switzerland: a retrospective cohort study. *BMC Geriatrics* 2015;15:114
- Kohler D, Widmer M, Weaver F. Les effets du nouveau financement hospitalier sur la qualité des prestations hospitalières stationnaires. Etude principale 1re étape 2008-2012. Obsan Rapport 62. Neuchâtel: Observatoire suisse de la santé, 2015
- Reich O, Signorell A, Busato A. Place of death and health care utilization for people in the last 6 months of life in Switzerland: a retrospective analysis using administrative data. *BMC Health Services Research* 2013;13:116
- Roth S, Pellegrini S. Le potentiel de transfert du stationnaire vers l'ambulatoire. Analyse pour une sélection d'interventions chirurgicales. Etude réalisée sur mandat de l'Office fédéral de la santé publique (OFSP). Obsan Dossier 63. Neuchâtel: Observatoire suisse de la santé, 2018
- Weaver F. Trajectoires sur dix ans des personnes atteintes de maladie chronique. Obsan Dossier 54. Neuchâtel: Obsan, 2016 (Kurzfassung: Weaver F. Längsschnittanalyse chronisch Kranker. Verlauf ihres Gesundheitszustandes, ihrer Inanspruchnahme von Leistungen der Gesundheitsversorgung und ihrer Arbeitsmarktbeteiligung. Obsan Bulletin 9/2016)
- Widmer M, Kohler D. Behandlungsketten und ihre Schnittstellen. Obsan Bulletin 1/2015. Neuchâtel: Schweizerisches Gesundheitsobservatorium
- Widmer M, Weaver F. Der Einfluss von APDRG auf Aufenthaltsdauer und Rehospitalisierungen. Auswirkungen von Fallpauschalen in Schweizer Spitälern zwischen 2001 und 2008. Obsan Bericht 49. Neuchâtel: Schweizerisches Gesundheitsobservatorium, 2011

Zellweger U, Bopp M. Record Linkage der Medizinischen Statistik 2000-2006 der Krankenhäuser mit der Swiss National Cohort (SNC). Swiss National Cohort Report Nr. 6. Forschung und Dokumentation 32, Institut für Sozial- und Präventivmedizin der Universität Zürich, 2010

Zellweger U, Bopp M. Record Linkage der Medizinischen Statistik der Krankenhäuser mit den Statistiken der natürlichen Bevölkerungsbewegung (BEVNAT) – eine Machbarkeitsanalyse. Swiss National Cohort Report Nr. 4. Forschung und Dokumentation 29, Institut für Sozial- und Präventivmedizin der Universität Zürich, 2006

Zellweger U, Bopp M. Record Linkage der Todesfälle in der Medizinischen Statistik der Krankenhäuser und der Statistik der Sozialmedizinischen Institutionen 2002-2012 mit der Swiss National Cohort (Swiss National Cohort Report Nr. 9). Forschung und Dokumentation 38, Institut für Epidemiologie, Biostatistik und Prävention der Universität Zürich, 2017

Zellweger U, Bopp M. Record Linkage der Todesfälle in der Statistik der Sozialmedizinischen Institutionen 2006-2008 mit der Swiss National Cohort. Forschung und Dokumentation 34, Institut für Sozial- und Präventivmedizin der Universität Zürich, 2011

Bestelladresse:

Institut für Epidemiologie, Biostatistik und Prävention der Universität Zürich
Hirschengraben 84, CH-8001 Zürich, Tel.: 044 634 46 11, e-mail: astrid.bruderer@uzh.ch

Schriftenreihe des Instituts für Epidemiologie, Biostatistik und Prävention (vormals: Sozial- und Präventivmedizin) der Universität Zürich:

Forschung und Dokumentation

- Nr. 1** Hornung R., Wydler H., Vontobel J., Tschopp A., Gutzwiller F.:
vergriffen Evaluation der Aidspräventionskampagne in Zürcher Schulen. 1989. (140 Seiten)
- Nr. 2** Gass R.:
Osteoporose: Prävention und soziale Kosten. 1990. (32 Seiten)
- Nr. 3** Jacob S., Gutzwiller F.:
Auswirkungen von Haferprodukten auf den menschlichen Serumcholesterinspiegel.
Ergebnisse einer Literaturrecherche. 1990. (51 Seiten)
- Nr. 4** Sempach R.:
vergriffen Prozessevaluation der gemeindenahen Prävention in Aarau. Ergebnisse einer qualitativen
Befragung von Verantwortlichen der Stiftung "AARAU eusi gsund Stadt" sowie
Personen aus dem politischen und sozio-medizinischen Umfeld. 1990. (48 Seiten)
- Nr. 5** Anliker P., Tschopp A., Gutzwiller F.:
Das Blutcholesterin und seine Bedeutung 1989. Resultate einer repräsentativen Ärzte-
befragung in der Schweiz. (57 Seiten)
- Nr. 6** Gass R.:
Auswirkungen der Arbeit an Bildschirmgeräten auf Schwangerschaft und Frucht. 1990.
(25 Seiten)
- Nr. 7** Arbeitsgruppe "Suchtprävention/Gesundheitsförderung":
vergriffen Suchtpräventionskonzept. Verfasst im Auftrag der Direktion des Gesundheitswesens des
Kantons Zürich. 1991. (84 Seiten)
- Nr. 8** Alvo, K., Hornung, R., Tschopp, A., Fuchs, W., Schaub, N.:
Intravenöser Drogenkonsum und Aidsprävention. Ergebnisse einer Befragung von 223
DrogenbenutzerInnen am Zürcher Platzspitz. 1991. (53 Seiten)
- Nr. 9** Hinnen U., Schüler G., Hotz P., Gutzwiller F.:
Hirntumoren bei der Anwohnerschaft der Deponie "Bärengaben" in Würenlingen.
Risiko-Beurteilung auf Grund von Häufigkeitsberechnungen des Zürcher Krebsregisters
und von Emissions-/Immissionsdaten. 1991/92. (29 Seiten)
- Nr. 10** Anliker P., Marti B., Gutzwiller F.:
Einstellungen und Verhaltensweisen der Schweizer Ärzte zu den kardiovaskulären
Risikofaktoren, insbesondere der Hypercholesterinämie. Ergebnisse einer Repräsen-
tativbefragung 1991. Verhaltensänderung der Ärzte von 1987 bis 1991 -Übersicht über
drei Befragungen. 1991. (47 Seiten)
- Nr. 11** Kerker M., Gutzwiller F.:
Messung der HIV-Seroprävalenz in der Schweizer Bevölkerung? Machbarkeitsstudie
einer anonymen HIV-Reihenuntersuchung auf freiwilliger Basis in einer Stichprobe der
Allgemeinbevölkerung. Entwicklung einer adäquaten Überwachungs-Strategie für die
HIV-Epidemie in der Schweiz. 1992. (70 Seiten)

- Nr. 12** Wellauer-Weber B., Brodmann P.:
Ernährungsinformation in der deutschen Schweiz. Erarbeiten einer Adressliste; Ziele, Aktivitäten und Ausbildung. Studienarbeit im Nachdiplomstudium "Humanernährung" an der ETH Zürich. 1991. (41 Seiten)
- Nr. 13** Gostynski M., Marti B.:
Stürze und ihre Konsequenzen bei institutionalisierten Betagten der Stadt Zürich: eine prospektiv - epidemiologische Studie. Resultate der ESBAZ - Studie. 1992. (90 Seiten)
- Nr. 14** Leuzinger A.D.:
Ernährungserziehung in Zürcher Kindergärten. Studienarbeit im Nachdiplomstudium "Humanernährung" an der ETH Zürich. 1993. (27 Seiten)
- Nr. 15** Vogel D.:
Anwendung und Erfolg von Diäten. Studienarbeit im Nachdiplomstudium "Humanernährung" an der ETH Zürich. 1993. (29 Seiten)
- Nr. 16** Pfister R.:
Anthropometrische Messungen. Theorie und praktische Anwendung. Studienarbeit im Nachdiplomstudium "Humanernährung" an der ETH Zürich. 1993. (43 Seiten)
- Nr. 17** Bally-Gerber E.:
Schneller postmenopausaler Knochenmassenverlust: ernährungs(mit)bedingt? Studienarbeit im Nachdiplomstudium "Humanernährung" an der ETH Zürich. 1993. (34 Seiten)
- Nr. 18** Bopp M., Schüler G.:
Neue Ansätze zur geographischen Mortalitätsanalyse. Nachweis langfristiger Krankheitshäufungen auch bei kleinen Fallzahlen. 1997. (146 Seiten)
- Nr. 19** Davatz F., Schmid M., Wydler H., Narring F., Michaud P.-A.:
Indikatoren unsafes Verhaltens bei Jugendlichen. 1997. Zusatzauswertung zum Projekt "Sexualité des adolescents et sida: processus et négociations autour des relations sexuelles et du choix de la contraception". Institut universitaire de médecine sociale et préventive, Lausanne 1997. (90 Seiten)
- Nr. 20** Gass R.:
vergriffen Welche Faktoren beeinflussen bei 20jährigen Frauen den täglichen Milchkonsum? (59 Seiten)
- Nr. 21** Bopp M.:
Record Linkage der anonymisierten Individualdaten aus Todesursachenstatistik und Volkszählung 1990. Machbarkeitsanalyse am Beispiel der Sterbefälle im Kanton Aargau 1990-1995. 1998. (50 Seiten)
- Nr. 22** Gretler S., Gutzwiller F., Hartmann G., Ronco C., Somaini B., Stutz Steiger Th.:
Prioritäre Gesundheitsbereiche in der Schweiz. Grundlagen für eine zukunftsgerichtete Gesundheitspolitik in der Schweiz auf der Basis einer Expertenumfrage 1998. (51 Seiten)
- Nr. 23** Bopp M., Gutzwiller F.:
Rekord Linkage von Volkszählung 1990 und Todesursachenstatistik 1990-97 in der Schweiz. Dokumentation zum Nationalfonds-Projekt 32-55884.98. 2000. (132 Seiten)

- Nr. 24** Hättich A., Bopp M.:
Evaluation möglicher Identifikationsvariablen für eine Verknüpfung der Individualdaten der Eidgenössischen Volkszählungen 1970, 1980, 1990 und 2000 (Swiss National Cohort Report Nr. 2). (86 Seiten)
- Nr. 25** Hämig O., Bopp M., Stähli R.:
Konzept zur Gesundheitsberichterstattung im Kanton Zürich 2003-2007. (71 Seiten)
- Nr. 26** Hättich A., Bopp M.:
Evaluation einer Verknüpfung der Ausländerstatistik PETRA mit den Volkszählungen 1980-2000 (Swiss National Cohort Report Nr. 3). 2006. (81 Seiten)
- Nr. 27** Zellweger U., Bopp M., Gutzwiller F.:
„Treffpunkt Herz“ – eine Präventionsaktion der Schweizerischen Herzstiftung: Herzinfarktrisiken sind in der Schweiz häufig. 2006. (30 Seiten)
- Nr. 28** Egger I., Reuter H., Bauer, G.:
Evaluation des Programms Allez Hop (2004-2006). Schlussbericht Juli 2006. (53 Seiten)
- Nr. 29** Zellweger U., Bopp M.:
Record Linkage der Medizinischen Statistik der Krankenhäuser mit den Statistiken der natürlichen Bevölkerungsbewegung (BEVNAT) – eine Machbarkeitsanalyse (Swiss National Cohort Report Nr. 4). 2006 (32 Seiten)
- Nr. 30** Bopp M., Gutzwiller F.:
Die Swiss National Cohort 1990-2000 – eine soziodemografische Datenbasis für longitudinale Gesundheitsanalysen (Swiss National Cohort Report Nr. 5). 2007. (59 Seiten)
- Nr. 31** Schmid M., Zellweger U., Gutzwiller F.:
Re-Evaluation der Rauchentwöhnungskurse der Fachstelle «nicht mehr rauchen» der Krebsliga des Kantons Zürich 2001-2005 Effektivität und Erfolgsfaktoren. (45 Seiten)
- Nr. 32** Zellweger U., Bopp M.:
Record Linkage der Medizinischen Statistik 2000-2006 der Krankenhäuser mit der SNC (Swiss National Cohort Report Nr. 6). 2010. (25 Seiten)
- Nr. 33** Tarnutzer S., Bopp M.:
Demografische Alterung und Krebshäufigkeit. Entwicklung der Zahl der jährlichen Krebsfälle nach Geschlecht und Lokalisation bis 2030 in der Schweiz und anderen entwickelten Ländern. 2011. (31 Seiten)
- Nr. 34** Zellweger U., Bopp M.:
Record Linkage der Todesfälle in der Statistik der Sozialmedizinischen Institutionen 2006-2008 mit der Swiss National Cohort (Swiss National Cohort Report Nr. 7). 2011. (17 Seiten)
- Nr. 35** Schmid M., Meidert U., Zellweger U., Gutzwiller F.:
Re-Evaluation der Rauchentwöhnungskurse «nicht mehr rauchen» der Krebsliga des Kantons Zürich 2006–2010 – Effektivität und Erfolgsfaktoren. 2012. (84 Seiten)
- Nr. 36** Tarnutzer S., Bopp M.:
Alter(n), Arbeit und Gesundheit. Eine illustrierte Charakterisierung der 40- bis 64-jährigen Bevölkerung der Schweiz. 2013. (78 Seiten)

- Nr. 37** Meyer J., Faeh D., Bopp M.:
Mortalität nach Berufsgruppen in der Schweiz 1990-2008: Machbarkeitsstudie und Empfehlungen (Swiss National Cohort Report Nr. 8). 2017. (63 Seiten)
- Nr. 38** Zellweger U., Bopp M.:
Record Linkage der Todesfälle 2002-12 in der Medizinischen Statistik der Krankenhäuser und der Statistik der Sozialmedizinischen Institutionen mit der Swiss National Cohort (Swiss National Cohort Report Nr. 9). 2017. (39 Seiten)
- Nr. 39** Mütsch M., Brunner S., Hämmig O., Puhan M.A.:
Psychische Gesundheit im Kanton Zürich: Welche Messgrößen wählen wir für die Jüngsten? Policy Brief and Policy Dialog vom Mai 2017. 2018. (59 Seiten)
- Nr. 40** Zellweger U., Bopp M.:
Identifikation und Bereinigung von Inkonsistenzen in der Medizinischen Statistik der Krankenhäuser (2000-16) und der Statistik der Sozialmedizinischen Institutionen (2007-16). 2019. (44 Seiten)